NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

THE EXPONENTIAL EXPANSION OF SIMULATION: HOW SIMULATION HAS GROWN AS A RESEARCH TOOL

by

Matthew J. Powers

September 2012

Thesis Advisor:	Susan M. Sanchez
Second Reader:	Thomas W. Lucas

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE September 2012	3. REPORT TYPE AND DATES COVERED Master's Thesis
4. TITLE AND SUBTITLE The Exponential Expansion of Simulation: How Simulation has Grown as a Research Tool		5. FUNDING NUMBERS
6. AUTHOR(S) Matthew J. Powers		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number: N/A.		
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		12b. DISTRIBUTION CODE A

13. ABSTRACT (maximum 200 words)

Simulation has overcome critical obstacles to become a valuable method for obtaining insights about the behavior of complex systems. George Box's famous assessment that "all models are wrong, some are useful" referred to statistical models, but should now be reimagined to reflect that many simulation models are "right enough" to aid in decision making for important practical problems. Over the past fifty years, simulation has transformed from its beginnings as a brute-force numerical integration method into an attractive and sophisticated option for decision makers. This is due, in part, to the exponential growth of computing power. Although other analytic approaches also benefit from this trend, keyword searches of several scholarly search engines reveal that the reliance on simulation is increasing more rapidly. A descriptive analysis paints a compelling picture: simulation is frequently a researcher's preferred method for supporting decision makers and may often be the "first resort" for complex real world issues.

14. SUBJECT TERMS Simulation, Data analysis, Design of Experiment, Simulation History			15. NUMBER OF PAGES 105
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
 Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

**THE EXPONENTIAL EXPANSION OF SIMULATION:
HOW SIMULATION HAS GROWN AS A RESEARCH TOOL**

Matthew J. Powers
Lieutenant Commander, United States Navy
B.S., United States Naval Academy, 2001

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

**NAVAL POSTGRADUATE SCHOOL
September 2012**

Author: Matthew J. Powers

Approved by: Susan M. Sanchez
                    Thesis Advisor

                    Thomas W. Lucas
                    Second Reader

                    Robert F. Dell
                    Chair, Department of Operations Research

iii

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Simulation has overcome critical obstacles to become a valuable method for obtaining insights about the behavior of complex systems. George Box's famous assessment that "all models are wrong, some are useful" referred to statistical models, but should now be reimagined to reflect that many simulation models are "right enough" to aid in decision making for important practical problems. Over the past fifty years, simulation has transformed from its beginnings as a brute-force numerical integration method into an attractive and sophisticated option for decision makers. This is due, in part, to the exponential growth of computing power. Although other analytic approaches also benefit from this trend, keyword searches of several scholarly search engines reveal that the reliance on simulation is increasing more rapidly. A descriptive analysis paints a compelling picture: simulation is frequently a researcher's preferred method for supporting decision makers and may often be the "first resort" for complex real world issues.

THIS PAGE INTENTIONALLY LEFT BLANK

THESIS DISCLAIMER

The data that were collected and analyzed in this thesis were obtained in large part with the use of the Naval Postgraduate School Library online scholarly search engines. Each search engine offered its own advantages and disadvantages that will be discussed throughout this paper. Such limitations included the inability to filter searchable items, such as reviews and letters to the editor, in order to separate them from the results that included scholarly articles. While this may slightly affect the results, it is important to note that the searches for the keywords "linear programming" and "optimization" were performed *exactly* the same as the search for "simulation" throughout all search engines. The behavior of the search was not conducted in such a way as to skew the results. Every effort was taken to ensure that there was no bias in favor of simulation that would have affected the manner in which I gathered the data.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ACM   Association for Computing Machinery

ACM DL   Association for Computing Machinery Digital Library

AIIE   American Institute of Industrial Engineers

AIS   Article Influence Score

CACM   Communications of the ACM

CI   Confidence Interval

DOE   Design of Experiments

EBSCO   Elton B. Stephens Company

ENIAC   Electronic Numerical Integrator and Computer

IBM   International Business Machines

ICBM   Intercontinental Ballistic Missile

GPL   General Programming Language

GSP   General Simulation Program

INFORMS Institute for Operations Research and Management Sciences

JOC   Journal on Computing

JSTOR   Journal Storage

LPAA   Lifetime Professional Achievement Award

LP   Linear Programming

LP-I/II   Laboratory Problem I/II

NOLH   Nearly Orthogonal Latin Hypercube

ORSA   Operations Research/Systems Analysis

PDES   Parallel Discrete Event Simulation

RAND   Research and Development

RNG   Random Number Generation

RVG   Random Variate Generation

SPL   Simulation Programming Language

WSC   Winter Simulation Conference

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

> Simulation by computation in engineering has increased its usefulness to such an extent that there are now at least a dozen large computers, several hundred medium size computers, and several thousand small computers at work on these problems in the United States.

> —Hurd (1954)

Computing capability and its impact on simulation impressed *some* scholars in the 1950s, but certainly not *all* scholars. The 1958 article, **"Simulation Techniques in Operations Research"** by John Harling described simulation as a "last resort." Harling considered simulation to be analogous to waving the white flag and settling for an approximate solution when it becomes too difficult to break a problem down into a "simple" model. However, three sentences after Harling called simulation a method of last resort, he went on to mention that "large machines are so often involved" when simulating that the cost often outweighs the benefit. This accusation seems laughable today, because Harling had yet to realize the phenomenon that Gordon Moore would formalize in 1965 when he predicted the exponential growth of computing power. Simulation was once described as a brute-force means of determining an answer. It would make sense to assume that the "brute-force" argument has since been dismissed, given modern computational capabilities (Harling, 1958).

Astonishingly, the argument that simulation should be a last resort is, unfortunately, alive and well fifty years later. In the 2008 article, "Dynamic Allocation of Airline Check-In Counters: A Queueing Optimization Approach," by Parlar and Sharafali, the authors constructed a complex queueing model to "optimize" the number of employees that should man the counters at a modern airline. Aside from a number of oversimplified, unrealistic assumptions (such as exponential service times, independence, the assumption that customers are queuing for a single aircraft, and a known number of only one customer at a time arrivals) that allowed for tidy math, the article claimed that the problem should be tackled via "analytical treatment" rather than "resorting" to simulation because that approach is "more realistic" and the "results are on firmer ground" (Parlar & Sharafali, 2008). By dismissing the use of simulation, Parlar and

Sharafali (2008) have placed themselves at odds with many researchers in operations research, as well as practitioners in many fields.

This thesis shows that reliance on simulation has grown exponentially, much like the computing power that supports such models. Since these models allow for stochastic parameters in place of wide assumptions, they account for the simple fact that you can't model everything. An analytical model may be an exact solution for the assumed parameters, but the solution is only as good as the question being studied. Simulation provides a solution for a question that is constantly changing. An analytical solution may be the exact answer to an approximate question.

This thesis is not designed to discredit analytical models. Nor is it designed to champion simulation. Such an effort would be a waste, as the results of this study show that more and more researchers have come to the conclusion that simulation is a valuable method. The growth of simulation is reflected in a growing number of scholastic journals, many of which are focused entirely on simulation. Furthermore, this growth is seen in an increase in simulation oriented articles in established scholastic journals such as *Operations Research, Computers and Operations Research, Management Science, INFORMS Journal on Computing,* and *Naval Research Logistics*.

The search engine utilized by *Naval Research Logistics* ("Torpedo") led to an interesting graphic and inspired search engine based keyword searches for "simulation" and related words. It was this portion of research that provided the most convincing and satisfying results. Through data collection via varying search engines, exploration of the increase in use of simulation was not limited to the scope of specific journals. The year-by-year keyword, title, and abstract searches by journal display a dramatic increase in the frequency and proportion of simulation-inspired articles. Each search engine allows for an "advanced search" that enables filtering of articles, year-by-year, in such a way that only those articles with the central topic of simulation appear on the results page. Aside from "Torpedo," the search engines provided by the Naval Post Graduate School library online database used for this research are JSTOR (Journal Storage), INFORMS, EBSCOhost (Elton B. Stephens Co.), and ACM DL (Association for Computing

Machinery Digital Library). These search engines also display the amount of results per year, so data collection became a simple (yet painstaking) matter of recording the amount of results and updating the year.

Prior to developing a graphic display of the data, it became apparent that the frequency of "simulation" (and related words) was growing rapidly over the years. However, when plotted, the data provided some striking results (see Figure S1).
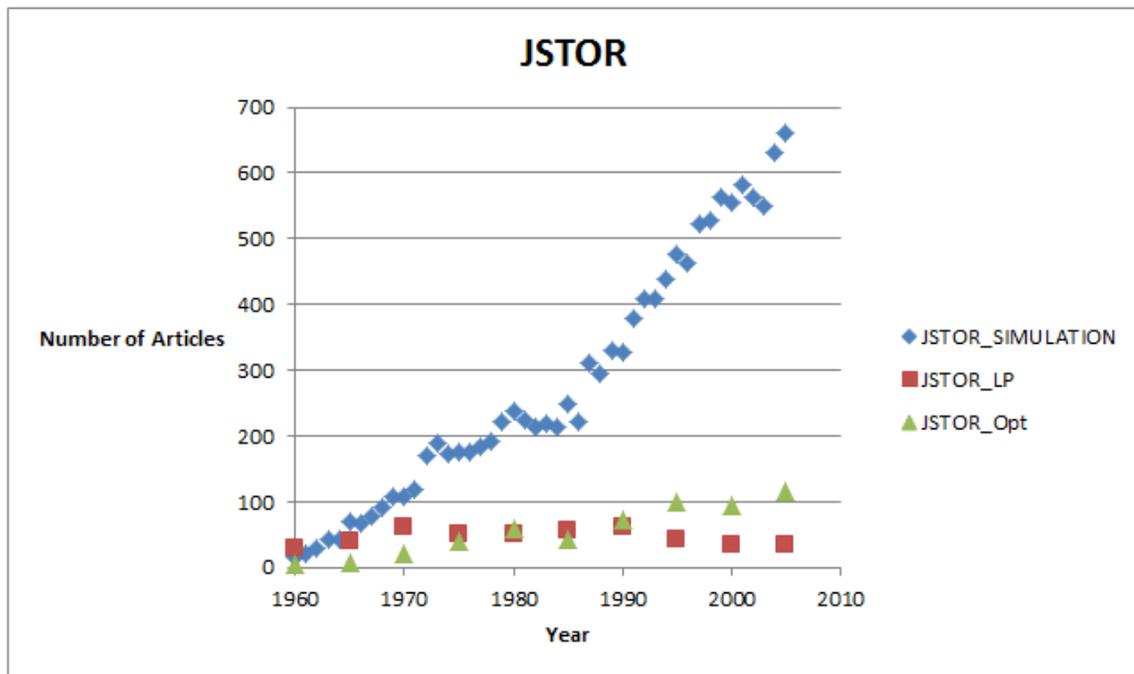
Figure S1. JSTOR data with "optimization" and "linear programming" results.

Notice that in Figure S1 that the blue simulation data is obviously growing much quicker than the "optimization" and "linear programming" data (also collected with related words). This figure is only a single example of the growth tendencies that this research reveals. The data set provides more than a convincing set of plots to display the significance of simulation. The data also allowed for in-depth regression analysis to model the specific nature of the growth and also prompted further study into which years experienced some particularly important leaps in the frequency of simulation use. A historical study of the scholarly journals and articles that coincide with these years tells

the story of the ever growing opinion that simulation is an important key to unlocking the mysteries of many of the most complex systems.

For a complex system to be modeled, it must be broken down and perhaps even simplified. If a system can be suitably simplified, it may still be worthwhile to solve analytically. That being said, analytical models are not gaining the popularity of stochastic models that have been simulation based. A comparison of search engine searches for analytical keywords "linear programming" and "optimization" showed that the growth of these particular topics has been much less dramatic over time.

A word that goes hand in hand with the idea of the passage of time is "advancement." Advancement refers to traversing a timeline as well as breaking beyond barriers. As barriers are broken in computing technology, so have they been broken in approaches to modeling the real world. It has become relatively simple to conduct 1,000,000 experiments on a simple model with random parameter. Thusly, the levels of complexity for models have increased dramatically. These abilities have intrigued many researchers and decision makers, and that fact has certainly inspired more and more researchers to consider simulation to be a method of first resort.

ACKNOWLEDGMENTS

the responsibilities of raising two boys who possess the stubborn energy of their mother and the charming tendency towards hijinks of their father, it would not have been nearly as fun! Daddy loves you, boys!

I hope that the influence I have on my children echoes the influence that my parents have had on me. I find myself inspired largely due to the fact that I consider how proud my mother would have been to hear of my accomplishments. My father has imbued in me the wonderful combination of work ethic and a tireless sense of humor. He has also exemplified the importance of perseverance in the face of tragedy, a virtue personified in the love he found in Mary Beth, and she herself personifies unfathomable love and support.

Support was not lacking in my military life or during my time at NPS. Among the great Americans who have made my list of "mentors" are Captain (USN) Jeff "Pancho" Davila, Commander (USN) David "Deke" Slayton, and most recently, Commander (USN) Harrison Schramm. Skipper Davila affirmed in me the necessity to keep my personality, regardless of my critics, because he "digs my gig" (first FITREP debrief). Deke Slayton has been a reassuring voice of reason and comfort throughout many professional and personal milestones and obstacles. Commander Schramm has been a shining example of the kind of passion that can exist in the career of an aviator even after he has said goodbye to the cockpit.

I'd like to give an extra special thanks to Dr. Thomas Lucas for quite literally saving the day when I approached him with a particularly important need for a thesis topic. If there is one example of a teacher and academic who can make contagious a passion for a topic (simulation), it is him. I consider myself very fortunate to have been able to study under such a brilliant mind. My fortune continued when I found myself in the presence of my advisor, Dr. Susan Sanchez. Her patience and keen understanding of the study of simulation has made the development of this thesis a smooth process. I am unable to express the gratitude I feel towards you, Professors Lucas and Sanchez, which is uncommon for me as I usually am unable to keep my mouth shut!

The efforts and enthusiasm of my advisers is evident in the fact that portions of this research have been summarized in the article by Powers, Sanchez, and Lucas (forthcoming), an article whose publication will result in a presentation given to the attendees of the 2012 Winter Simulation Conference in Berlin, Germany.

Last, but certainly not least, I would like to express my sincere gratitude to Dr. Lee Schruben for providing me with his 2011 letter nominating Peter Lewis for the INFORMS Simulation Society's Lifetime Professional Achievement Award. Also, thank you to Dr. David Kelton for providing useful data, articles and insight that were extremely useful in the development of this thesis. Thank you both gentlemen for invaluable words of encouragement!

My short time at Naval Postgraduate School will prove to be the most valuable and memorable years of my life! I experienced new adventures and levels of happiness with my amazing family. I developed friendships and bonds that have already reached well beyond the classroom. I have come to understand what it means to think like an analyst, and someday I will be proud to call myself one!

THIS PAGE INTENTIONALLY LEFT BLANK

I.   INTRODUCTION

> An object of operations research in business is to apply scientific methodology to the solution of management problems; in particular, the operations research tool, simulation, is useful.
>
> —Hurd (1954)

A.   BACKGROUND AND LITERATURE REVIEW

Simulation had gained a reputation as being useful as a tool in operations research long before the astonishing growth in computing power. In order to fully understand and appreciate the impressive rise in the use of simulation, one must pay respect to the genesis of its expansion. In the early computer days of the 1950s and 1960s, simulation was by no means a new concept. The use of simulation certainly predates computers, with perhaps the most famous example being Buffon's estimation of π in 1777 (Nance & Sargent, 2002). A later example that predated computers was in 1908 when William Sealy Gosset manually simulated the probability density function for his "Student's 't' Distribution" (Fisher, 1925). However, researchers in the 1950s and 1960s were certainly beginning to adopt the opinion that simulation could be used as a means of putting years of research and data to work. Similarly, the use of simulation motivated further research and stimulated advancement across all operational fields. The one-two punch of research-inspired simulation and simulation-inspired research sparked the acceptance of simulation's significance in management, industry, and military operations, just to name a few important fields.

1.   Early Research Benefits of Simulation

In his article, "The Simulation of a Large-Scale Military Activity," Murray A. Geisler described a pair of experiments that he named "Laboratory Problem I" and "Laboratory Problem II" (LP-I and LP-II), whose respective goals were to model the complex organization known as the United States Air Force, and to model the non-existent intercontinental ballistic missile (ICBM) force. The objectives of these experiments were similar in that they both were meant to serve the Air Force, but they

differed in that LP-I was based on existing research, while LP-II would require further research to design a system that lived only in the minds of the researchers (Geisler, 1959).

a. *Simulation Motivated by Research*

In 1953, the Air Force and the RAND (Research and Development) Corporation set forth to design and implement models to aid in analyzing decision making variables such as "characteristics of demand for spare parts, including their demand probability distributions, developing inventory decision rules, determining cost parameters, time lags, depletion penalties, etc." After years of toying with "difficult and time-consuming" mathematical formulations, the Air Force and RAND were "dissatisfied" and no-doubt frustrated with their lack of progress in creating satisfying solutions to their problems. Eventually, the decision was made to use a "detailed man-machine simulation" to "bridge the gap between research and implementation" (Geisler, 1959).

Simulation allowed for the desired level of detail to be incorporated into a decision making tool that benefited from years of research. Simply put, the Air Force could take advantage of policies and data that existed in other organizations and absorb them into its own set of rules. Given the complexity of a real-world organization such as the Air Force, the knowledge that could be gained from the years of data at their fingertips could only be optimized through computer-based simulation. LP-I was an early example of the wide-reaching advantages of simulation, and researchers were so enthused at the results that they extended their experimentation in order to model the future; the yet to be realized ICBM force (Geisler, 1959).

b. *Research Motivated By Simulation*

At the time of Geisler's writing, the second experiment (LP-II) was still under way. The idea behind LP-II was to create a model that would represent what would come to be known as the ICBM force. Such a task would require further research, and the data collected during said research would be immediately implemented into the model. While the model would eventually serve as the foundation for an extremely

ambitious project, at the time of Geisler's article, it was serving as pure motivation for the kind of research necessary to develop such a significant system. Often times in experimentation, one simply does not know which questions need to be asked. LP-II was an example of simulation helping to generate those important questions.

Geisler described two examples, LP-I and LP-II, of computer based simulation aiding a particular branch of the military. The military, however, is not alone in its early recognition of this revolutionary tool as a breakthrough in research and development. Expensive decisions are routinely made in the worlds of management and industry. As computing capabilities improved, simulation was starting to gain acceptance as a means of helping with those decisions.

c.   *Recognition in Management, Industry, and Military*

In business and other problems where simulation is being used today, it is likely that we are at the stage where only simulation can be used to attack many of the important problems.

—Teichroew (1965)

In the business world, the dollar is the bottom line. With that philosophy in mind, it makes sense that simulation was gaining popularity in management and industry as the cost of sophisticated computers was falling. In 1954, a time coincident with the Air Force and RAND experiments described earlier, Cuthbert C. Hurd wrote of a machine-shop scheduling problem that turned to simulation when other methods failed. Much like those involved in the design of LP-I, researchers from the 1953 Operations Research Society of America meeting attempted to solve scheduling problems through analytical means. Particularly, they tried to apply "the theory of game and of linear programming," but found these methods to be of little help (Hurd, 1954). What *was* of great help was the IBM (International Business Machines) 701, a machine that "calculated directly" the solution to the machine-scheduling problem through simulation (see Figure 1).

Figure 1.   IBM 701 (From IBM archives, www.IBM.com, n.d.)

Some of Hurd's words were unique in that he complimented mathematicians and their deep exploration of complex systems and the formulations that motivate them. He credited their studies for benefitting the physical world and for igniting the desire for further research. Hurd's tip of the hat to theoretical mathematicians is in stark contrast to the common simulation versus analytical solution arguments that exist among scholars. Hurd suggested that "simulation by computation" was the key to unearthing deeper mysteries. He went on to mention that the method of simulation was continuing to gain popularity thanks to the falling "unit cost of carrying out computation." In other words, computers were getting cheaper. Thinking back to the "bottom line" mentality, a cheap machine that computed the cheapest decision rules made perfect sense as the weapon of choice for high-power decision makers (Hurd, 1954).

Speaking of weapons, let us recognize the early uses of simulation as a tool in strategic-level decision making. That tool is known as the "war game." In the experiment described by Geisler, the Air Force resorted to simulation to piece together financial and management type puzzles, but the military has also been able to take advantage of computational power to gain an edge on the field of battle. In a 1955 paper entitled, "Simulation as an Aid in Model Building," R. P. Rich discussed a model that

was based on a fleet air defense system. He stressed that the fact that a military example was being used was purely "coincidental," because almost any complex scenario could have been used to demonstrate the utility of simulation. Rich's paper is a valuable resource in that it illustrates many benefits that branch from simulation, and those benefits shall be discussed later, but his words on war gaming are particularly relevant in the way they highlight the way that simulation has been used to model battlefield operations (Rich, 1955).

In a peculiar way, Rich used the word "machine" to describe simulation before he or the world had fully come to realize the impact that machines would have on the subject. When he wrote that, "a war game is just a special type of simulator, with the opposing teams so many parts of the machine," he was not referring to a computer that ran the model, but rather the individuals whose actions provided a stochastic element of the scenario. The decisions made by the people playing the game were the moving "parts," that is, they were a random component. Rich did, however, mention a "simple device" that played a role in "emulation" for the fleet air defense scenario. It was the combination of the device and the opposing team "parts" that provided the basis for this very special type of simulator. The war game would prove to be a type of model that would play an important role in strategic and operational development across all branches of the military (Rich, 1955).

We have briefly discussed the early research benefits of simulation and the reach of those benefits across the fields of management, industry, and military decision making. The next section concerns some of the objectives of simulation.

2.   Objectives of Simulation

By far, the early work in simulation and that which has been dominant in management science and operations research over the history is system analysis, where the intent is to mimic behavior to understand or improve system performance.

—Nance and Sargent (2002)

In the previous section, it was made clear that simulation can be used as an aid to research or, perhaps even more useful, as a means of inspiring research. This section

discusses some objectives of simulation beyond (and including) research. An understanding of these objectives should help to explain why simulation became such a hot topic during the age when computers began to find their place in the world. It is particularly interesting to notice that the objectives of simulation are strikingly similar to the objectives of analytical methods of computation such as linear programming or game theory. The fact that some who championed those closed-form analytical methods considered simulation to be a less sophisticated, "brute-force" means of computation will be discussed later. For now, let us focus on the objectives of simulation, which were defined by Richard E. Nance and Robert G. Sargent to be system analysis, education and training, acquisitions and system acceptance, and of course, research (Nance & Sargent, 2002).

a. System Analysis

Simulation is used with the purpose of system analysis when the "intent is to mimic behavior to understand or improve system performance" (Nance & Sargent, 2002). Such was the objective in Geisler's presentation of LP-I, where the "system" being mimicked was the entire United States Air Force (Geisler, 1959). Rich used similar words to express the same objective when he described a "device which simulates the physical situation to be analyzed," where the "situation" was the United States Air Force fleet air defense system. Rich went on to stress that the "first purpose" of the simulator was "to provide a nucleus for collecting data" (Rich, 1955). The point of referring back to previously discussed authors and their arguments is to demonstrate the point that system analysis is a sort of umbrella objective that acts as the root of so many experiments. This analysis then branches out to satisfy the other objectives.

b. Education and Training

With thorough analysis, simulation can provide the kind of data, numerical or otherwise, to provide the kind of insight that can be used for education and training. The spirit of this education objective is exemplified in Rich's paper where he specifically mentioned that the air defense simulator was "also helpful in explaining the results of the analysis to the people for whom they are intended" (Rich, 1955). As for training, a

simulation can be useful assuming it is a valid model of real life in that it can prepare an operator for expected responses of the system to inputs. A specific example is that of an aircraft simulator. A pilot or aircrew member can become familiar with aircraft responses when he or she introduces full-opposite rudder and control inputs. With enough trials, the pilot can then familiarize himself or herself with the proper procedures to prevent the aircraft from falling out of the sky.

If familiarization is to be considered a special form of education, then simulation is one of the most valuable educational tools available. Rich emphasized the way in which the simulator at the center of his study "led naturally and directly to a mathematical formulation of the problem" (Rich, 1955). Such a claim counters the "brute-force…last resort" arguments against simulation (Harling, 1958). Furthermore, Rich mentioned that the researchers behind the fleet air defense simulation became "educated" about particular details of the system such as the "probability of death as a function of time" which had "received very little attention until the simulator emphasized its importance" (Rich, 1955).

c. Acquisition and System Acceptance

System acceptance is a logical by-product of system analysis. Through simulation, one has the ability to determine if a system has what it takes, so to speak, to live up to predetermined standards. Among the varied uses of simulation, Rich stressed the importance of being able to provide the "same answer" to "the people who are to make use of it" (Rich, 1955). What he meant is that simulation is a means of answering the question, "will this system do what is needed?"

The above question mirrors the sentiment of the questions presented by Nance and Sargent, where they describe the acquisition objective of simulation as the answer to, "Does the system meet the requirement?" or, "Does a subsystem contribute significantly to the improvement of the larger system performance" (Nance & Sargent, 2002).

d.   Research

As "research" was a central idea in the first section, it will only be expanded on here, specifically in the way that research leads to an enhancement of system analysis and therefore all the other objectives. Daniel Teichroew wrote in his article, "A History of Distribution Sampling Prior to the Era of the Computer and its Relevance to Simulation" that, "in simulation as in distribution sampling, it is to be hoped that more emphasis will be placed on the methodological problems associated with its use, and that the output of the simulation experiments will be used to stimulate and guide theoretical analysis" (Teichroew, 1965). These words perfectly embody the relationship between research and analysis, and they hint at the future of simulation and the need for improvement of the "methodological problems associated with its use." While he does not explicitly credit "past mathematicians" for their contributions to the physical world, Teichroew did describe simulation as a bridge between real world systems and the theory behind them, and his words suggested just how the two opposing forces actually benefit one another. Rich further solidified the connection between research and analysis when he recognized that the thought which "frequently comes first to mind" when speaking of simulation is its use as "a computing device to provide numerical answers" (Rich, 1955).

Numerical answers are the output from the very analytical "nucleus for collecting data" that is simulation (Rich, 1955). In that way, system analysis begets research, and research enhances system analysis. Similarly, as already mentioned, simulation benefits the theoretical world of analysis, and has been shown to lead directly to mathematical formulation. With that in mind, it makes sense that by the 1960s, simulation was poised to take off as a revolutionary means of research, with benefits that far outweigh the costs.

> [Simulation] can serve as a focus in the collection of data, as an aid in the construction of a mathematical model, and as a useful bridge between operations and research. The cost of such simulators in time, money, and manpower is so small, and their advantages in the early stages of study are so great, that they appear to deserve much fuller exploitation and discussion than they have enjoyed in the past.
>
> —Rich (1955)

3. The Launch of Computer Simulation

> The use of simulation precedes computers, either analog or digital. Described by some authors as "artificial sampling," a manual Monte Carlo method was employed by Buffon to estimate π in a study documented in 1777.
>
> —Nance and Sargent (2002)

The focus of this section is on the early years of simulation in the computing era, as opposed to an entire history of computer simulation. With early research benefits of simulation and some objectives having had already been discussed, a brief explanation of computers and their role in simulation will serve as a platform from which to launch what will eventually become an analysis of the growth of simulation as a research tool. This section covers the first use of computer simulation, followed by some discussion of the subject throughout the 1960s, and finally, the "scholarly disrespect" that simulation still contends with despite its increasingly valuable contributions.

a. First Computer Use (WWII)

While simulation "in the case of continuous Monte Carlo methods" had been in existence well before the 1940s, computer simulation began during World War II (Nance & Sargent, 2002). Specifically, ENIAC (Electronic Numerical Integrator and Computer) was employed during the Manhattan project in 1945 in order to "solve an important problem in nuclear science" (Fritz, 1994). To fully appreciate the growth in simulation use as it relates to the growth in computing capability, one must have a clear picture of what a top-of-the-line computer looked like (see Figure 2).
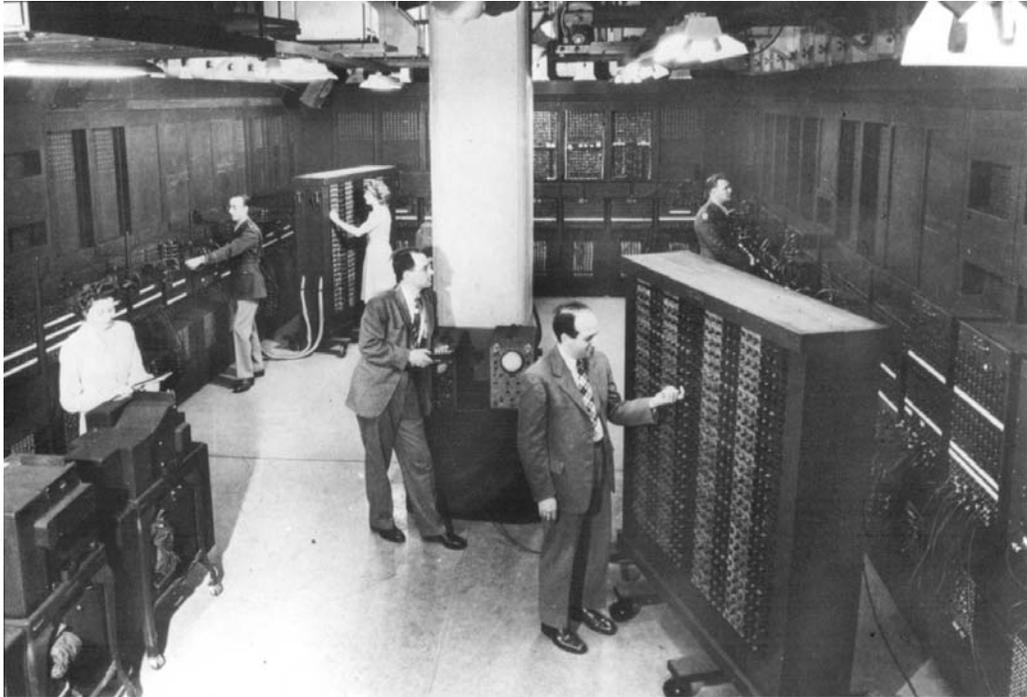
Figure 2.   The ENIAC (From The Franklin Institute, www.fi.edu, n.d.)

While the computing power available during WWII pales in comparison to what is common today, computing speed of the ENIAC was essentially 1,000 times faster than any computer available at the time (Fritz, 1994). In order to be operated, ENIAC's individual units had to be connected by cables and sequencing was determined by setting switches (Fritz, 1994). Such a description paints an image of something less like a "computer" and more like the "man-machine" device mentioned in Geisler's article on simulation (Geisler, 1959). Either way, the method of simulation had moved beyond dropping needles onto the floor and the world was set to witness an evolution in computing technology that would resemble ideas that were conceived only in science-fiction.

b.   *1950s through 1960s*

Computers in the 1950s and the 1960s were a long way from even the simplest hand-held devices that are common today, but they were formative in the history of simulation. For instance, Simulation Programming Languages (SPLs) became more commonplace, which offered a translation from the General Programming Language

(GPL) used by many computers so that the modeler did not have to rely on general programming libraries. With a suitable language in place, the next steps were the improved representational capability and "functional library organization needs" that would lead to the GSP (General Simulation Program) of Tocher and Owen in 1960 (Nance & Sargent, 2002). Clearly, programming that was geared towards simulation was being developed, but there remained problems at the heart of the theory behind computer simulation.

Random number generation (RNG) and random variate generation (RVG) for SPL programming relied on a simple technique which could not have remained true to randomness properties (Nance & Sargent, 2002). At best, computers at the time could produce "pseudorandom numbers" that would not place a "heavy demand on the machine's main store" (Harling, 1958). However, whether the random generation relied on published tables or an indeterminate process, the method was flawed and did not live up to the usual statistical demands. For instance, the multiple replications required for output analysis were limited by the RNG/RVG capabilities, and oftentimes the analysis was simply an "estimation of the mean without concern for variance estimation" (Nance & Sargent, 2002). The mechanical analogue randomizer was one of the simple yet crude methods of random generation. This particular machine was used to study the behavior of queues (see Figure 3).
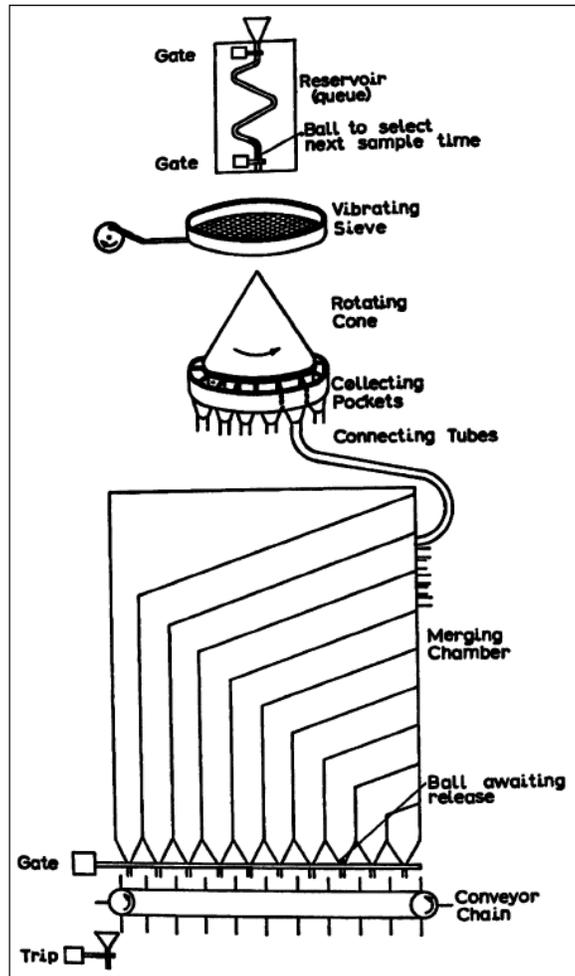
Figure 3.   United Steels' mechanical analogue randomizer (From Hollocks, 2006).

It was evident that there existed some fundamental theoretical problems behind computer simulation. Those flaws could lead to misleading solutions, and those solutions were not quickly available for inspection.

Robert G. Sargent recalled that the turnaround time for a graduate student's submission at the University of Michigan in the 1960s could be as long as one to two days. This slow and cumbersome method, while an improvement over the ENIAC days, was due in large part to the slow reading of the 80 column IBM cards (see Figure 4) required for programming (Nance & Sargent, 2002).
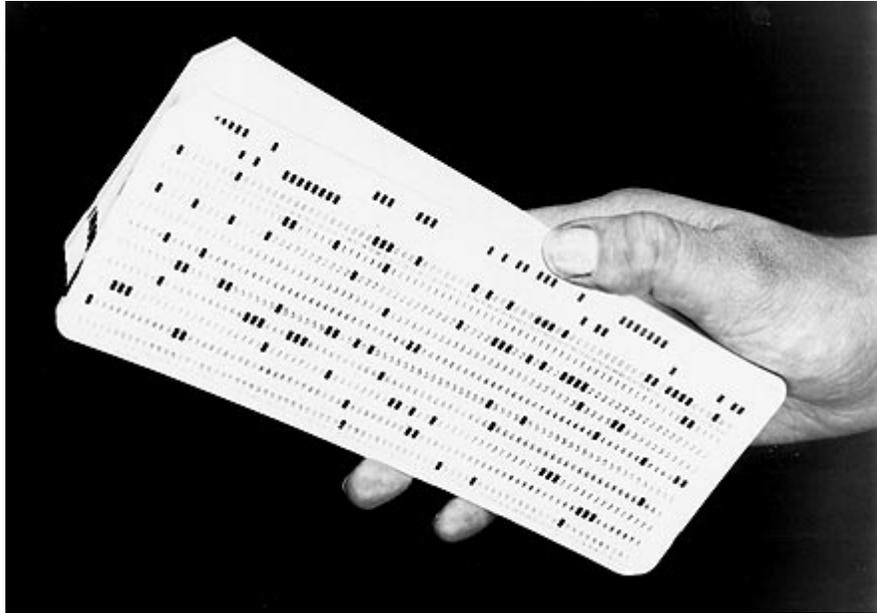
Figure 4.   80 column IBM card (From IBM archives, www.IBM.com, n.d.)

        The IBM card represented "programming" in the early 20th century until more sophisticated methods were born. The 80 columns defined the 80 characters per card, and the cards were read at approximately 133 characters per second. Not only did this primitive form of simulation programming require days of turnaround time to obtain a solution, but the debugging and analysis steps were often tedious.

        Despite the drawbacks, computer simulation was catching on as a popular method for arriving at a solution to a complex problem by the end of the 1960s. As with any hot topic, computer simulation was met with its share of critics, who were happy to dismiss this method as a fad. R. W. Conway's 1963 article, "Some Tactical Problems in Digital Simulation," perfectly expressed the mixed feelings that were developing towards simulation.

> The use of a digital computer to perform simulated experimentation on a numerical model of a complex system is an increasingly important technique in many disciplines today. It has made possible the systematic study of problems when analytic or solvable models are not available and actual "in situ" experimentation is impractical. Although the brief history of this form of simulation is generally favorable, there remains much room

for improvement in the practice of the technique. I believe that, in general, simulation models take longer to construct, require much more computer time to execute, and yield much less information than their authors expected

—Conway (1963)

One of the most telling phrases in the above paragraph is, "the study of problems when analytic or solvable models are impractical," in that it hints at the idea of relying on simulation when other methods are not possible. An attitude that was prevalent in the 1960s is also captured in the final sentence, when Conway all but dismisses simulation because of the large amount of computer time to execute the model. In his argument, Conway used many words to criticize simulation when he could have used just two: last resort. The phrase "last resort" was to be found among many scholarly articles in the 1950s and 1960s (Conway, 1963).

It has been often said that a simulation is a last resort.

—Harling (1958)

Statisticians found that distribution sampling is a technique that is better than nothing but is less desirable than analytical techniques… Therefore, as in statistical problems, the technique should be used only as a last resort.

—Teichroew (1965)

Most operations research analysts look upon digital computer simulation as a "method of last resort."

—Wagner (1969)

Many professionals in management science and operations research cast simulation as the "method of last resort" and expressed the view that "anyone could do it."

—Nance and Sargent (2002)

### c.	*Scholarly Disrespect*

> Until the late-1960s…fundamental developments in simulation were readily accepted in the scientific literature. About this time an attitude of scholarly disrespect seemed to emerge.

> —Nance and Sargent (2002)

As with any skill, the ability to develop and execute a valid and successful simulation was frowned upon by many who were experts in alternate abilities. It had been argued that simulating a problem was analogous to finding the answer without ever fully understanding the question, and for that reason "anyone could do it" (Nance & Sargent, 2002). Furthermore, the time and effort required to "solve" a complex problem was not worth obtaining the solution. Imagine the disrespectful attitude adopted by some scholars as that of a student that must prepare for an exam. Said student could fully immerse into the text, the notes, and the theory behind the material, or simply gather a shallow collection of knowledge that is the minimum required to pass the exam. The former strategy not only guarantees a passing grade, but it cements a firm grasp of the material at hand, thus fulfilling the purpose of examination as a means of student evaluation. The latter strategy satisfies the letter grade criteria, but not the spirit of education. The "anyone can do it" attitude of analytic scholars echoed the idea that the solution obtained via simulation represented a shallow collection of knowledge (Harling, 1958). Astonishingly, the argument that simulation should be a last resort is, unfortunately, alive and well fifty years later. In their 2008 article, "Dynamic Allocation of Airline Check-In Counters: A Queueing Optimization Approach," Parlar and Sharafali constructed a complex queueing model to "optimize" the number of employees that should man the counters at a modern airline. Aside from a number of oversimplified, unrealistic assumptions that allowed for tidy math, the article claimed that the problem could be tackled via "analytical treatment," rather than "resorting" to simulation (Parlar & Sharafali, 2008). By dismissing the use of simulation, Parlar and Sharafali (2008) have placed themselves at odds with many researchers in operations research, as well as practitioners in many fields.

It has been said that no press is bad press. With so much debate over the validity of computer simulation as a research tool, the topic was gaining much "press" by the 1960s. Many of the debates focused on the "misuse of simulation," combined with a preoccupation with "mathematical sophistication" that was not present in simulation (Nance & Sargent, 2002). Whatever the center of the argument, computer simulation was being discussed, particularly in the form of scholarly articles in respected journals. This coverage, paralleled with impressive technological advances, allowed for the topic of computer simulation to shrug off any disrespect and gain strides in reputation.

4. Computer Simulation Acceptance and Coverage

The emergence of simulation departments (or areas) in the archival journals in the late 1970s gave evidence of a reputation regained. An increasing number of simulation researchers were finding outlets for quality publications.

—Nance and Sargent (2002)

a. Publications and Dates

Nance and Sargent (2002) provided an extensive list of significant journals and publications that were born in the wake of computer simulation.

- *American Institution of Industrial Engineers (AIIE) Transactions* in 1976, with Nance as editor.
- *Management Science* in 1978.
- *Operations Research* in 1978, with Nance as editor.
- *Communications of Association for Computing Machinery (CACM)* in 1980, with Sargent as editor.
- *Operations Research/Systems Analysis (ORSA,* now *INFORMS*) *Journal on Computing* in 1989, containing an area on simulation with Nance as editor.
- *Transactions of the Society for Computer Simulation* in 1984.
- *Transactions on Modeling and Computer Simulation (TOMACS)* in 1990, with Nance as editor-in-chief.
- *Simulation Digest* from 1988 to 1994, the first online publication devoted explicitly to simulation interests.

A search of the NPS online eJournal database returned a list of publications and the years that they were established (see Table 1).

Title	Year Established
Simulation & gaming	1970
Simuletter	1971
Mathematics and computers in simulation	1977
Modelling and simulation in materials science and engineering	1992
Simulation practice and theory	1993
Building Energy Simulation User News	1996
Computational geosciences	1997
Journal of statistical computation and simulation	1997
JASSS	1998
Macromolecular theory and simulations	1998
International journal of simulation: systems, science & technology	2000
AISB journal	2001
Communications in statistics. Simulation and computation	2001
IJIGS. International journal of intelligent games & simulation	2002
Simulation modelling practice and theory	2002
Training and Simulation Journal	2002
IEEE International Workshop on Distributed Simulation & Real-Time Applications (DSRT)	2004
Internet Journal of Medical Simulation and Technology, The	2005
World Journal of Modelling and Simulation	2005
International journal of modelling & simulation	2006
Journal of simulation	2006

Table 1.   Simulation inspired publications and the years that they were established (From Naval Postgraduate School, http://www.nps.edu/Library/).

b.   *Journal Significance*

Not only were the scholarly outputs for computer simulation vast, but they have enjoyed significance in the academic world. A quick analysis of computer simulation inspired journals and publications shows a high level of significance based on some well-established metrics. An analysis of some of that significance is discussed in Chapter II of this thesis.

The remainder of this thesis shows how simulation has grown from the early computer years (starting with the year 1960) where it experienced much scrutiny through the year 2010. The analysis is not based upon scores or values, but rather through an extensive search of how often "simulation" and related words are in the title or are keywords in scholarly articles over the aforementioned years. The analysis begins with a database of frequencies and of simulation-centric articles in some particular journals. After inspection of several specific journals, this thesis analyzes a database

created by searching for "simulation" and related words once again, but through different search engines over the same span of years. The results of the search-engine data analysis display exactly what this thesis intends: Simulation as a research tool has grown exponentially (or at least rapidly) over time.

This growth, while not following the exact same trend as described in publication and conversation by Gordon Moore for computing power in 1965, owes itself to the impressive advances in computer technology and the usefulness of the method for solving complex, real world problems. This thesis does not intend to discuss computer history over time. It will, however, pinpoint some particularly influential years in the data analysis and investigate some possible reasons that certain years may have been so important. The idea behind such an investigation is that the significant years in the history of simulation will coincide with significant advances in computing technology *and* with relevant academic theory. With each advance, the arguments that existed against computer simulation in the early computer days fade. Simulation has evolved into a method of computation as sophisticated as the computer itself, and often is the method of first resort!

> It is true that a simulator which is to serve as a computer should have a high-speed capability…

> —Rich (1954)

B.   RESEARCH QUESTIONS

This thesis uses operations research methods, specifically data analysis techniques, to examine the frequency of occurrence of key words in various and specific Operations Research focused journal articles. Throughout the research, I pay close attention to the evolution of simulation and track the changing opinions across the operations research field. Specific questions include:

- (a)   How has simulation grown as a tool used by researchers?
- (b)   How does that growth compare to the growth of computing power as suggested by Moore's Law?
- (c)   How does that growth compare across journals?
- (d)   How does that growth compare across the search engines available on the NPS library eJournal database?

(e)   How does simulation growth across scholarly journals compare to the growth of keywords relating to examples of analytical methods such as optimization and linear programming?

(f)   Does analysis of the data indicate particularly influential years in the growth of simulation?

C.   BENEFIT OF THE STUDY

This research formally credits the methods of simulation modeling as a primary means of problem solving. While simulation, just as George Box said of all models, would most likely still be considered to be "wrong," this research shows that there exist many scholars that consider them to be "useful."

Simulation modeling has been utilized across such fields as the Department of Defense, industry, and beyond. This research proves that simulation has gained recognition as a valuable tool over the years to claim its spot as one of the invaluable means of not just finding the answer, but in asking a better question.

D.   THESIS ORGANIZATION

The remainder of this thesis is organized into the following chapters. Chapter II provides a quick analysis of the frequency of simulation-based articles in specific journals. Chapter III consists of a series of in-depth analyses of different search engines over the years where "simulation" (and related words) were supplied as the keywords. This chapter includes a description of the applied research method for the first search engine. The analysis includes a model for each search engine and its associated data set as well as a look at which years were highlighted as unusual and influential. These years may offer hints as to which events in computing history and academia had the most significant impact on the growth of simulation. The research is concluded in Chapter IV, where analysis of the research is summarized. The conclusions also revisit the comparison of simulation's growth as a research tool next to methods such as optimization and linear programming. This thesis ends with recommendations for further research. Of particular interest would be a deeper study into which events in computing and academic history did in fact have the most influence on simulation. Such a study

would allow for an understanding of which types of advancements play key roles, and perhaps inspire further benefits so as to optimize future use of simulation.

II. JOURNAL-SPECIFIC KEYWORD ANALYSIS

> The idea of numerical simulation is easy to understand, leading to its wide acceptance in diverse application areas spanning the physical sciences, engineering, and management.
>
> —Kelton (1994)

Chapter II begins with analysis of the scholarly significance of some of the journals mentioned in Chapter I. Note that these journals are not all the same as those that will be analyzed via the keyword search method. The significance analysis can be accomplished through a search of the NPS Library's link to the Web of Knowledge. By investigating the "Science Citation Reports" for the journals of interest, valuable metrics and graphics can be gathered for evaluation of a publication's scholarly significance (from http://admin-apps.webofknowledge.com.libproxy.nps.edu/JCR/). Among the metrics, each journal is assigned an "Article Influence" score (AIS). The website provides an explanation for the score. Another valuable evaluation tool is the "Journal Impact Factor," which, simply put, measures the number of times that articles from the journal have been cited.

The Article Influence determines the average influence of a journal's articles over the first five years after publication. The mean Article Influence Score is 1.00. A score greater than 1.00 indicates that an average article in the journal has above-average influence. A score of less than 1.00 indicates that an average article in the journal has below-average influence. The journal impact factor is a measure of the frequency with which the "average article" in a journal has been cited in a particular year. The impact factor will help you evaluate a journal's relative importance, especially when you compare it to others in the same field (http://admin-apps.webofknowledge.com.libproxy. nps.edu/JCR/).

The website also displays an "Impact Factor Trend Graph." This graph displays the Impact factor scores over a number of years, so one can envision any changes in significance over the years for a particular publication. The journals mentioned in the

first part of this section have displayed some impressive metrics and graphics in terms of their significance (see Figures 5 – 8).

Management Science (started in 1978)

Article Influence Score: 2.504 (recall that the mean AIS is 1.00, meaning that a score of 2.504 is well above average influence).
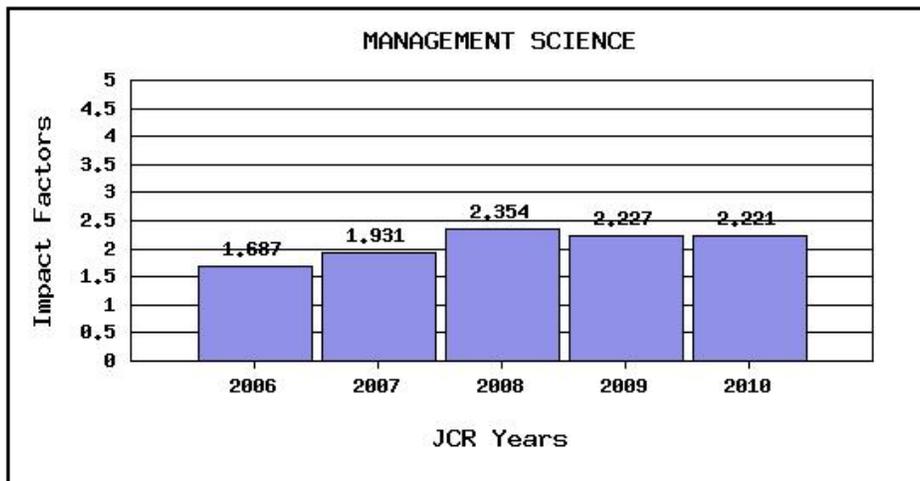
Figure 5.   Impact Factors for *Management Science.*

The Impact Factor mitigates the importance of absolute citation frequencies. It tends to discount the advantage of large journals over small journals because large journals produce a larger body of citable literature. For the same reason, it tends to discount the advantage of frequently issued journals over less frequently issued ones and of older journals over newer ones. Because the journal impact factor offsets the advantages of size and age, it is a valuable tool for journal evaluation (from http://admin-apps.webofknowledge.com.libproxy.nps.edu/JCR/).

Operations Research (started in 1978)

Article Influence Score: 1.928

OPERATIONS RESEARCH
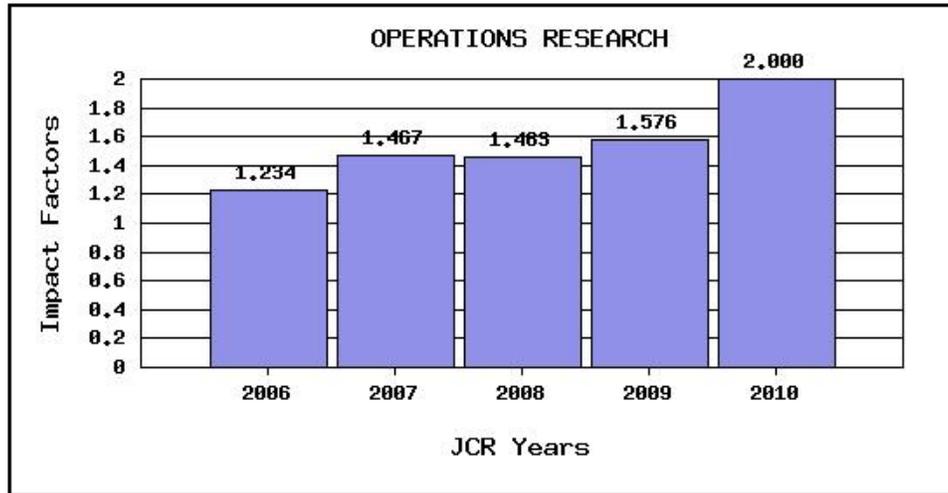
Figure 6.   Impact Factors for *Operations Research.*

Communications of ACM (started in 1980)

Article Influence Score: 0.829

COMMUNICATIONS OF THE ACM
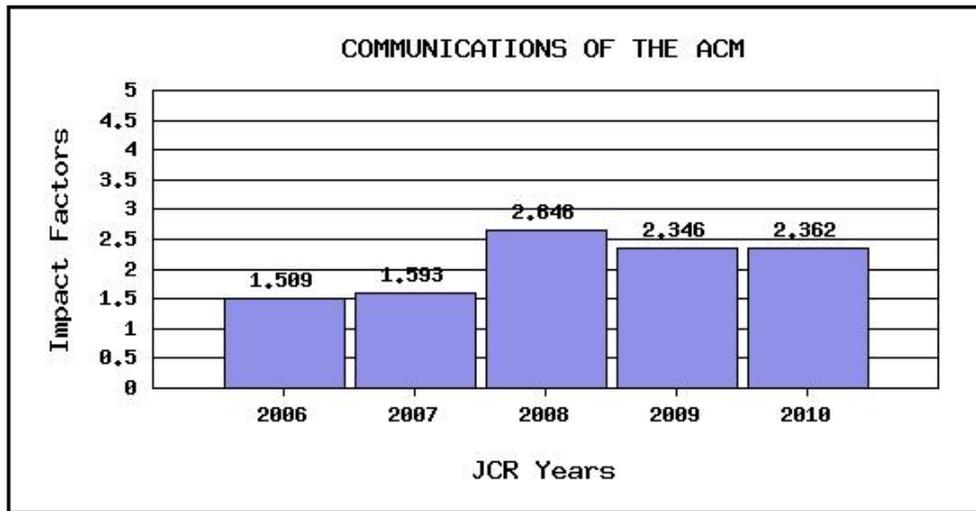
Figure 7.   Impact Factors for *Communications of the ACM.*

While CACM's AIS is slightly below average, the Impact Factors are on par with the other journals that have been discussed.

23

***INFORMS Journal on Computing* (started in 1989)**

Article Influence Score: 0.903

INFORMS JOURNAL ON COMPUTING

Impact Factors

| 0.865 | 0.907 | 1.041 | 1.318 | 1.172 |

2006 2007 2008 2009 2010

JCR Years
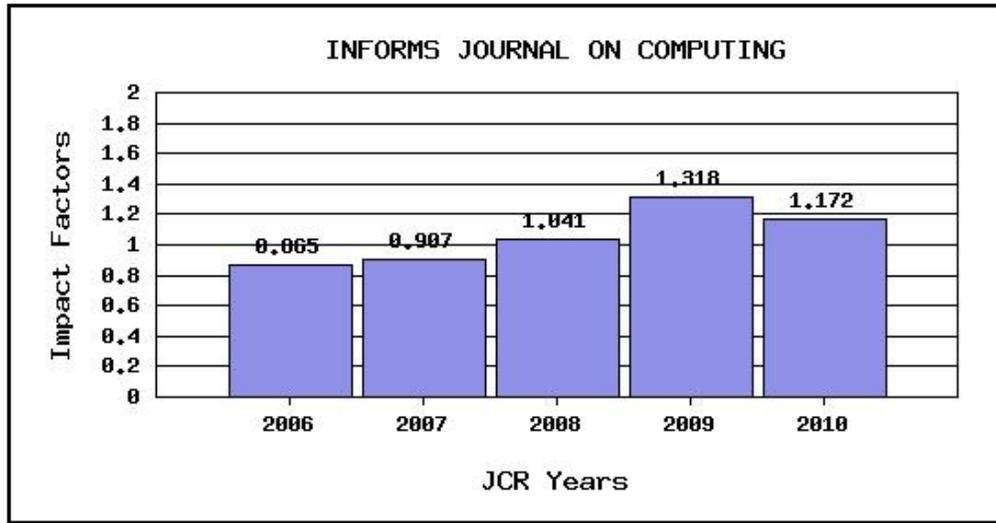
Figure 8.   Impact Factors for *INFORMS Journal on Computing.*

INFORMS JOC Impact Factors are lower than others, but have a higher AIS than CACM. It is important to keep in mind that the *INFORMS JOC* is not an entirely simulation-based publication; rather it dedicates a section to simulation.

What follows in this chapter is the keyword search analysis of several scholarly journals. The data represents the frequency of simulation articles by year across those publications. When fit to a model, it is possible to show that the positive slope of some of the curves is statistically significant. In this section, the words "Frequency" and "Year" represent the response and independent variables, respectively. In these models, those slopes are not particularly impressive, and hypothesis testing shows that there is not enough evidence to support the validity of some of those slopes. While the trends fail to impress, the models allow for some possible insight into what may have been happening behind the scenes in the OR community in its attitude towards simulation. This analysis is unique in that it relies on the old adage that a picture is worth a thousand words. With the ability to simply compare the data to its corresponding plot it is easy to extract the most influential years with respect to simulation in the history of a specific publication. Similarly, it is possible to gain some insight into the hot topics of the time. The bottom

24

line for this analysis is that the results are underwhelming as far as the growth of simulation is concerned. Underwhelming, that is, until the final analysis of a flawed search-engine (Torpedo Ultra) displayed a compelling graphic that hinted at a more appropriate way to conduct this research.

A.   OPERATIONS RESEARCH

An essential requirement for papers that are published in the INFORMS journal *Operations Research* is that they are "broad." That is, *Operations Research* is by no means a simulation-only publication. A look at how often simulation articles appear in such a publication should give a general feel as to the growth of simulation, especially if the results plotted over time display an uphill trend (see Figure 9).
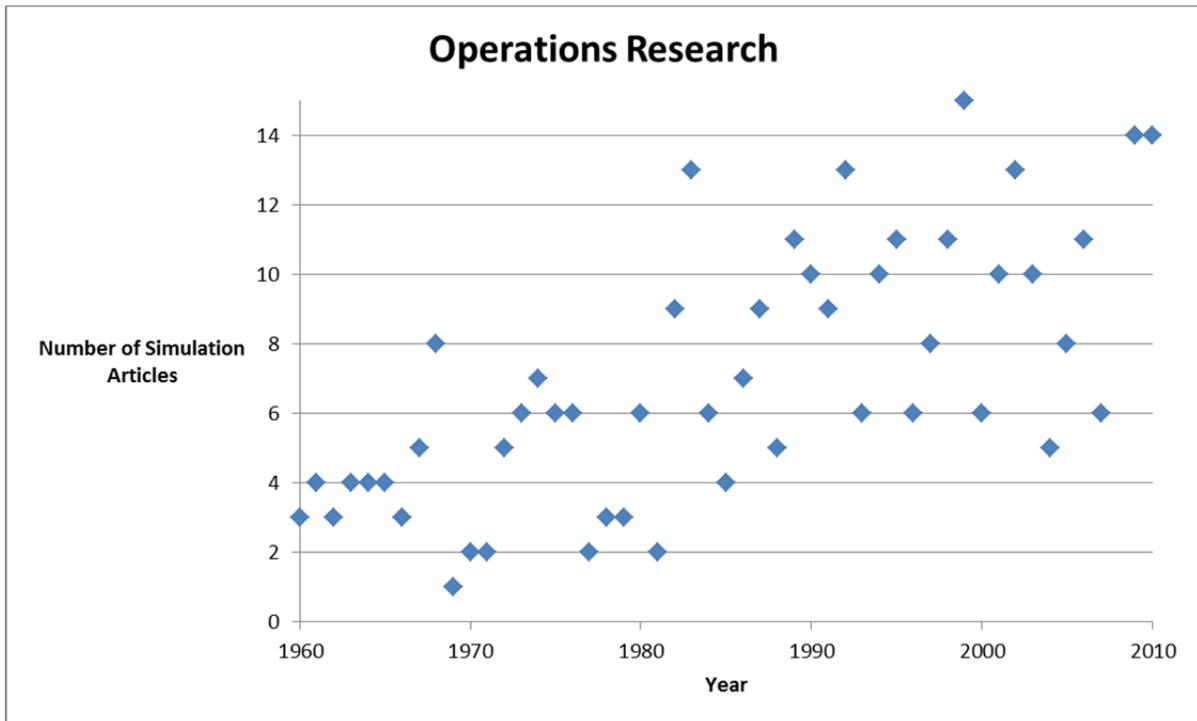

Figure 9.   *Operations Research* simulation articles over time.

For the data represented in Figure 9, we can fit a linear regression model with the frequency of articles regressed against the year. As will be the case in most of the models, the year 1960 is known as year 1, 1961 is year 2, and so on until 2010 which is

year 51. Such a fit is useful in performing a hypothesis test where the null hypothesis says that the slope of the line equals zero. Using the statistical software R (www.r-project.org), the result is the following model:

- Frequency = 2.26275 + 0.19050 * Year

This model has a p-value of < 0.0001 for the "Year" coefficient, which demonstrates the statistical significance of the uphill slope. The R^2 value of 0.4935 proves the fact that time is not solely responsible for the tremendous variation in this model. Factors that are nearly impossible to quantify, such as editor preferences, special issues, and the unknowable pace of research breakthroughs, may explain some of the variability, and a close look at the data itself will help to identify some peculiarities.

A peculiar year in this simple linear fit is 1989 (11 articles) as it is the first time in the data set where the number of articles doubles from the previous year to a relatively high (double-digit) amount. The common theme in some of the articles that were published in 1989 is statistical analysis through simulation. For instance, two articles dealt with sensitivity analysis by means of simulating a model ("Restricted Subset Selection Procedures for Simulation" by Sullivan and Wilson and "Sensitivity Analysis for Simulations via Likelihood Ratios" by Reiman and Weiss). These two, along with the article "Using Parallel Iteration for Approximate Analysis of a Multiple Server Queueing System" by Birge and Pollock, argue the statistical validity of simulation models. Birge and Pollock (1989) addressed the errors induced by unrealistic independence assumptions in Markov chain models. Their model provides an exact solution if the "assumption of independence among servers is valid," but their model is also useful in the more practical setting where such independence cannot be safely assumed. These three articles serve as examples in the history of *Operations Research* when simulation was finding its home in the statistical and analytical world (Birge & Pollock, 1989).

B.   COMPUTERS AND OPERATIONS RESEARCH

This publication is not an INFORMS journal. Nonetheless, it is interesting to see how simulation has found itself relevant in a publication whose title alone tells the reader

that the articles within will deal with the role that computers play in operations research. As such, the increase in simulation articles over time is more impressive than in *Operations Research*, yet testing shows that the growth remains linear, not exponential. A linear regression allows for a hypothesis test to statistically show that the slope of the line is greater than zero. However, the positive value of the slope coefficient can be ascertained with a glance at a plot of the data (Figure 10).
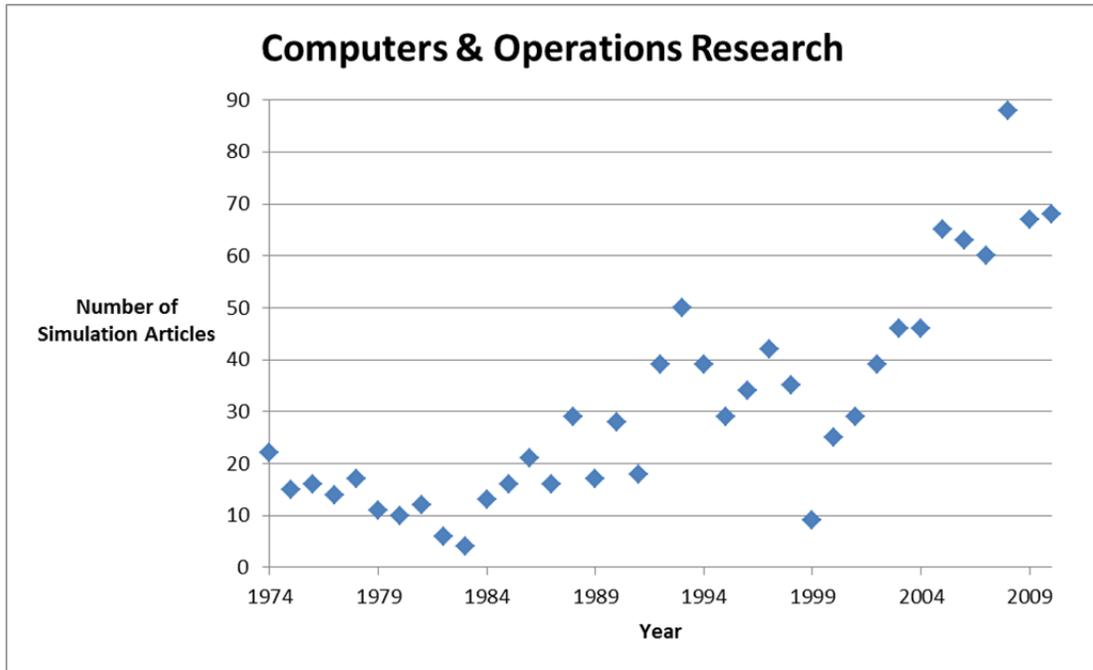
Figure 10.   *Computers & Operations Research* simulation articles over time.

Since the earliest available data for this publication is 1974, that is labeled as year 1 and 2010 is year 37. The data fits rather nicely into the following model:

- Frequency = 1.4234 + 1.5723 * Year

With a p-value much less than 0.0001 for the "Year" coefficient, it is safe to say that the null hypothesis is rejected and the slope is greater than zero. The R^2 value for this model is 0.6723, once again confirming the need for other factors to explain the variability. Heteroscedasticity does not appear to be a major problem. The significant

drop of simulation articles in the year 1999 could be viewed as an outlier. However, a closer look at the data suggests the possibility of a better fit by using a piece-wise linear regression.

The results drop from 35 simulation articles in 1998 to an astonishingly low nine articles in 1999. Perhaps the large drop in interest reflects the OR community's lack of astonishment with advancement within the field. As previously suggested, perhaps the drop is a result of editor interest. Interestingly, one of the nine articles, "Netsim: Java-based simulation for the World Wide Web," dealt with an object-oriented simulation package (Netsim) that utilized Java and its interoperability with the world-wide web. The "object-oriented structure of Netsim greatly facilitates" the ability to expand this general-purpose simulation language, and the world-wide network for which it was designed expounds the expansion (Veith, Kobza, & Kelling, 1999). It is unlikely that it is a coincidence that this publication realized a steady increase in simulation articles after 1999 as software came online that allowed for world-wide network access.

To test this theory, it is necessary to split the data into two sets and model each. The first set will be 1974 through 1998, with 1974 retaining its label as year 1. The second set will include 1999 through 2010 with 1999 assuming the label of year 1. By comparing the slopes of these two sub-models, it is possible to recognize that the behavior of the data set is significantly different after the year 1999, and perhaps that difference is owed to the advances mentioned in the previous paragraph. This experiment results in the following models:

- 1974 – 1998: Frequency = 6.3200 + 1.2154 * Year, with a 0.95 CI of (0.7365, 1.6943) on β_{Year}.

- 1999 – 2010: Frequency = 13.7121 + 5.6469 * Year, with a 0.95 CI of (3.8900, 7.4037) on β_{Year}.

The slope of the latter-years model is greater than four times the slope of the model of the early years. The fact that there is no overlap of either of the 0.95 confidence intervals (CI) on each slope coefficient confirms the statistically significant difference of

28

each slope. A hypothesis test on the slope of both models confirms statistically that the slopes are both greater than zero, but the R^2 value of the early model is 0.5451 while that of the latter model is 0.8368. This says that time accounts for more of the variability in the latter-years fit, and that it is unlikely to be a coincidence that this publication experienced a rising trend in simulation articles during those years.

C.   MANAGEMENT SCIENCE

Articles published in *Management Science* are primarily focused on the disciplines of economics, mathematics, psychology, sociology, and statistics.   Once again, this multidisciplinary publication does not deal specifically with simulation. The plot shows that the number of simulation articles varies over time, and the variation is rather consistent. Also, a frequent occurrence of large jumps and drops in simulation articles between years called for many windows of time in which to investigate (Figure 11).
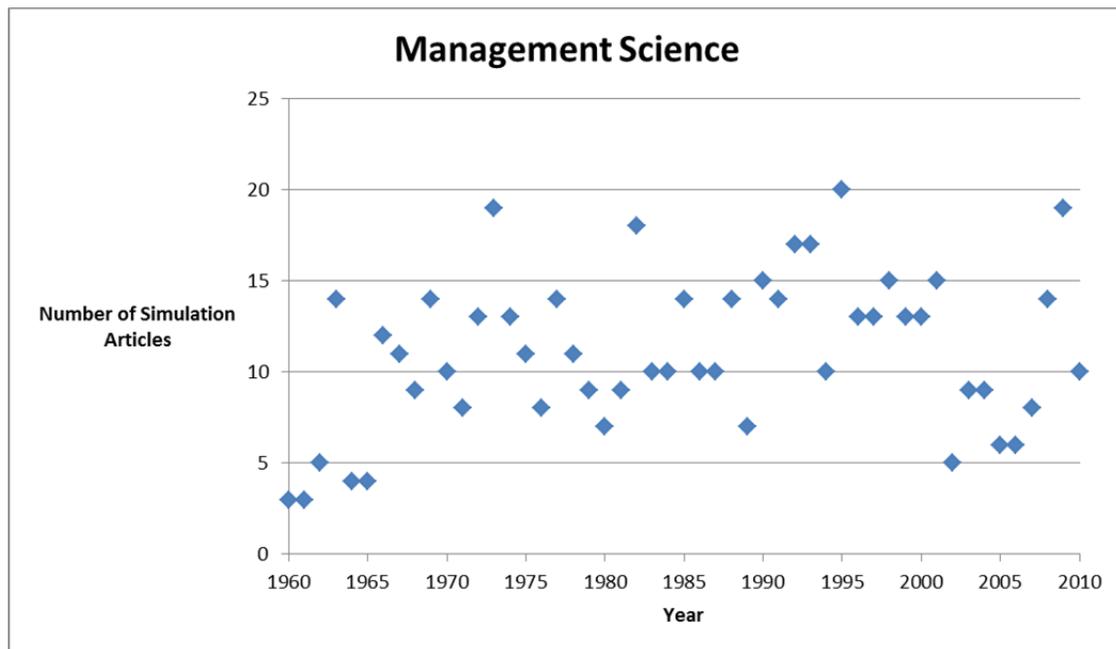

Figure 11.   *Management Science* simulation articles over time.

The model that results from this data set exemplifies the underwhelming nature of this portion of the research. The model is as follows:

- Frequency = 9.0455 + 0.0782 * Year

The p-value for the "Year" coefficient is 0.0517, meaning that there is not enough statistical evidence to reject the null hypothesis that the slope is equal to zero at the 0.05 level. That is, the *Management Science* simulation articles did not tend to rise significantly over the years. Furthermore, the R^2 value of 0.0782 not only expresses the inadequacy of time as a sole regressor in this model, but it makes it difficult to pinpoint influential years.

While it is not easy to see any influential years from the plot or from the regression, the jump in frequency from 1965 (4) to 1966 (12) led to an investigation of those years. Of particular interest was an article by K. D. Tocher (1965) dealing with the advancement of simulation languages. Tocher summarized some of the important features that were in development with tables (see Tables 2 – 4).

			TABLE 1		
Language	Basic language	Dominant type of entities	Principal authors	Machine implemented for	Reference
GPSS	None	Material	Gordon	IBM 7090	7, 8, 9
SIMPAC	SCAT (assembly)	Material	Lackner	IBM 7090	10, 11
SIMSCRIPT	FORTRAN	Material	Markowitz	IBM 7090	12
SIMULA	ALGOL	Material and machines	Nygaard and Dahl	Univac 1107	13
CSL	FORTRAN	Machines	Buxton, Laski	IBM 7090 IBM 1620 IBM 1410	14, 15
			Clementson	Honeywell 400/1400 200/2200	
ESP	ALGOL	Machines	Williams	Elliott 503	16, 17
GSP†	None	Machines	Tocher	Ferranti Pegasus Elliott 503‡	3, 6
MONTECODE	AUTOCODE	Machines	Buxton and Head	Ferranti Pegasus	18, 19
SIMON	ALGOL	Machines	Hills	Elliott 503 and an 8000 word 803	4

† GSP II—GSP I is now obsolete.
‡ In preparation.

Table 2. Tocher displayed the simulation languages that have already been developed and some that were being developed (From Tocher, 1965).

Language	Uniform distribution	Histogram	General continuous function	Method of fitting	Normal	Method	Negative exponential	Number of random number streams
GPSS	Yes	Yes	Yes	Piecewise linear	No	—	No	Indefinite
SIMPAC	Yes	Yes			Yes		No	1
SIMSCRIPT	Yes	Yes	Yes	Piecewise linear	No	—	No	1
SIMULA	No details yet released							
CSL	Yes	Yes	No	—	Yes	Sum of 10 random	Yes	Indefinite
ESP	Yes	Yes	No	—	Yes	Sum of 10 random	Yes	Indefinite
GSP	Yes	Yes	Yes	Modified polynomial	Yes	Special case of general	Yes	Indefinite
MONTECODE	Yes	Yes	No	—	Yes	Table look-up	No	1
SIMON	Yes	Yes	Yes	Not stated	No	—	Yes, special case of Erlang	10

Table 3.   Tocher's "Table 6" summarized the "facilities for sampling" offered by each language at the time. The advancement of random sampling was a hot topic among critics of simulation, and that topic was being addressed (From Tocher, 1965).

Language	Histogram forming facilities	Limit on number of classes	Calculation of mean	Calculation of standard deviation	Limit on number of distributions
GPSS	Yes	100	Yes	Yes	100
SIMPAC	Yes	None	Yes	Yes	None
SIMSCRIPT	No	None	Yes	Yes	None
SIMULA	No details yet released				
CSL	Yes	None	Not automatic but additional statements aid in its calculation	Not automatic but additional statements aid in its calculation	None
ESP	Yes	None	Yes	Yes	None
GSP	Yes	None	Yes	Yes	None
MONTECODE	Yes	14	Yes	Yes	Preset parameter
SIMON	Yes	None	Yes	Yes	None

Table 4.   Tocher's "Table 7" summarized the statistical analysis capabilities of the simulation languages available in 1965 (From Tocher, 1965).

31

The next big jump was from 1981 (nine articles) to 1982 (18 articles), and that jump was immediately followed by a drop to 10 articles in 1983. One of the 1982 articles featured in *Management Science* was "A Four-Moments Alternative to Simulation for a Class of Stochastic Management Models," by John F. Kottas and Hon-Shiang Lau. It seemed appropriate that this article existed in a year that would act as a local peak for the publication, as if to say "not so fast" to the popularity of simulation. These authors offered "a computational alternative to simulation for a large class of stochastic management models involving functions of random variables." Much like many articles that have served as critiques of simulation, this paper proposed a closed-form method of analyzing stochastic models. The article attempted to conclude with a bang, with its suggestion that "there are, of course, situations in which the [objective value] function is too complicated for the computational approach. In that case, simulation becomes the last resort." (Kottas & Lau, 1982). (*Management Science* simulation articles experienced a significant drop in 1983.)

Management Science simulation articles experienced their highest peak in 1995 with 20 articles. While no obvious similarity exists among these articles, the peak inspires a look at some of the happenings in the simulation software world as well as a flourish of interest in simulating optimization algorithms to arrive at a "determined objective function." A summary of seven of the top US commercial simulation software developers in 1990 showed the types of advancements that were being made (or planned) at the time (see Figure 12). Such efforts reflect the popularity of simulation in the commercial world. That is, simulation had broken far beyond the academic world, and companies were cashing in (Hollocks, 2006).

Better dynamic graphics or interface	5 packages
Improved ease of use	4 packages
Portability/transfer to more hardware	4 packages
Scheduling tools	3 packages
Expert Systems features	2 packages
Integration facilities to other software	2 packages
Business graphics	1 package*
Output analysis	1 package*
Experimentation features	1 package*
*nb: not the same package.	

Figure 12.   A review of seven leading U.S. simulation software developers' plans in 1990. (From Hollocks, 2006).

The "scheduling tools" and "experimentation features" listed above may have foreshadowed the emergence of an academic interest in optimization by simulation. This interest has "blossomed in the form of optimization packages" and was reflected in a pair of articles featured in *Management Science* in 1994. "Stochastic Optimization by Simulation: Numerical Experiments with the M/M/1 Queue in Steady-State" and its companion piece "Stochastic Optimization by Simulation: Convergence Proofs for the GI/G/1 Queue in Steady-State" described models that converged to a specified objective value without the use of closed form analytics. As a counter to the "last resort" argument, authors Pierre L'Ecuyer and Peter W. Glynn made it clear that simulation had become the only "viable" method to solve "complex (realistic) stochastic models" (L'Ecuyer & Glynn, 1994). Simulation articles dropped from 20 in 1995 to 13 in 1996. This was not a drop to below average, but rather a drop back to the mean value.

There was a significant drop from 15 articles in 2001 to five in 2002. (2002 was the final year with Professor Hau L. Lee as Editor-in-Chief of *Management Science*. While the drop may be a coincidence, it is interesting to note that Prof. Lee's background is in supply chain management and he has given many talks and has published many

articles dealing in optimization.) The "leveling out" of the simulation articles in recent years will be seen again in the search-engine analysis, but the leveling will be occurring on seemingly exponential curves. This phenomenon will be discussed later.

D.   INFORMS JOC

The *INFORMS Journal on Computing (JOC)* publishes original research articles on subjects that are at the "intersection of operations research and computer science." Simulation finds itself comfortably suited for that intersection. However, this publication is far from a simulation-centric journal. Pages in this publication are awarded to a variety of topics, and that variety is apparent in the plot of simulation articles over time, which displays yet another moderate growing trend with relatively equal variance above and below the mean (Figure 13).
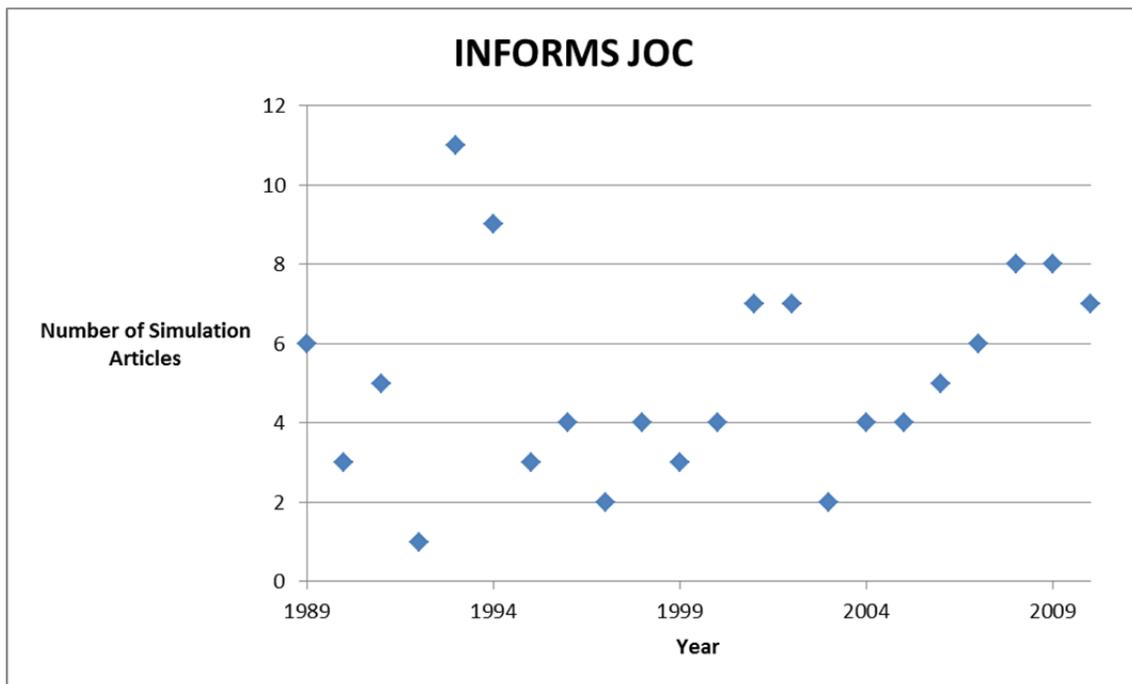

Figure 13.   *INFORMS JOC* simulation articles over time.

This model was fit with 1989 labeled as year 1. The resultant model is as follows:

- Frequency $= 4.2338 + 0.0785 *$ Year

Once again, this is an underwhelming model. The p-value for the "Year" coefficient is 0.3690, meaning that the coefficient is in no way significant. It would appear that there is no uphill trend in this data. Also, the R^2 value is extremely low at a value of 0.0785. Once again, the model does a poor job of explaining the variance, and the "significant" years are as easily ascertained by glancing at the data as they would be through sophisticated methods.

The 11 articles in 1993 and the nine articles in 1994 represent the most unusual years in this linear fit. Results jump from one article in 1992 to 11 articles in 1993. It turns out that seven of those 11 simulation articles were commentaries on Parallel Discrete Event Simulation (PDES). This serves as an example of special topics (in this case, PDES) accounting for some of the variability in the model. Any given PDES article, be it for or against the subject, discussed the (at the time) all too real problem of system performance when it came to running parallel simulation models. One common observation was the difficulty of gaining "acceptable speedup for parallel simulation in a network of work stations due to long communication delays" (Lin, 1993). The use of high-speed networks, which were certainly not as commonplace in 1993 as they are today, was suggested as a future means of dealing with the problem. It is safe to say that, as of this research nearly 20 years later, unacceptable network speed is much less of an issue.

The year 1994 remained high with nine simulation articles covering diverse topics that included simulation in networking and optimization. David Kelton authored a 1994 article, "Perspectives on Simulation Research and Practice," that discussed various subjects such as RNG, software, and new design techniques. Kelton provided a fantastic summary of the Winter Simulation Conference (WSC) attendance that included a graphic that, much like those that will be presented in this research, served as a compelling argument for the growth of simulation (see Figure 14).
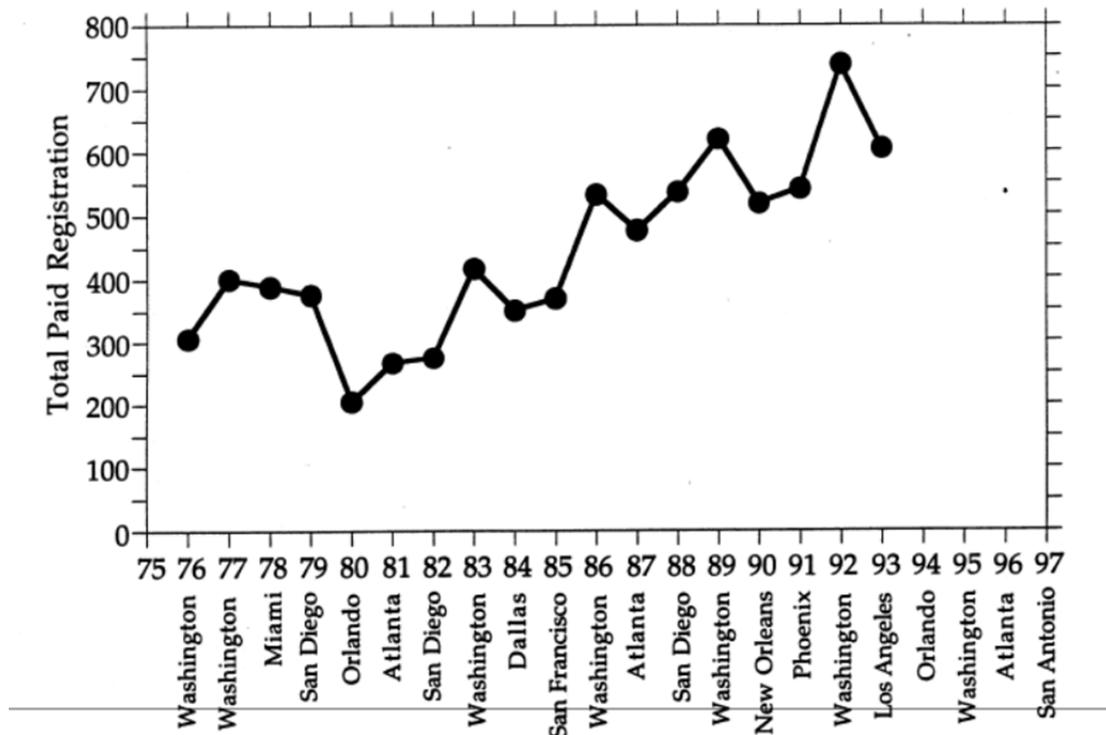
Figure 14.   WSC Paid Attendance from 1976 to 1993 (From Kelton, 1994).

Kelton also gave a nod to recent developments and the flourish of discussions concerning PDEs, which was apparent in the relatively high results for 1993. It would seem as if Kelton's paper has since become widespread, as the words in his concluding remarks concerning the "impact, influence and success" of simulation and the need for further research of "appropriate software" appear to foreshadow things that have come true (Kelton, 1994). There have been overwhelming advances in software, and design of experiments (DOE) is a growing field in the simulation discipline (and it will be discussed in greater detail within this research).

E.   NAVAL RESEARCH LOGISTICS

The publication *Naval Research Logistics* utilizes the "Torpedo Ultra" search-engine to search its archive. The search engine is, in a word, difficult. There is no obvious way to isolate keyword searches within specific publications. Various methods were attempted for this research and they resulted in error messages. However, the

frustration stemming from the attempt to gather data from Torpedo Ultra resulted in an influential graphic that displayed a substantial growth in simulation articles that existed in the Torpedo Ultra database (see Figure 15). Even if though these results do not exhibit exponential growth, they inspired a different search method.
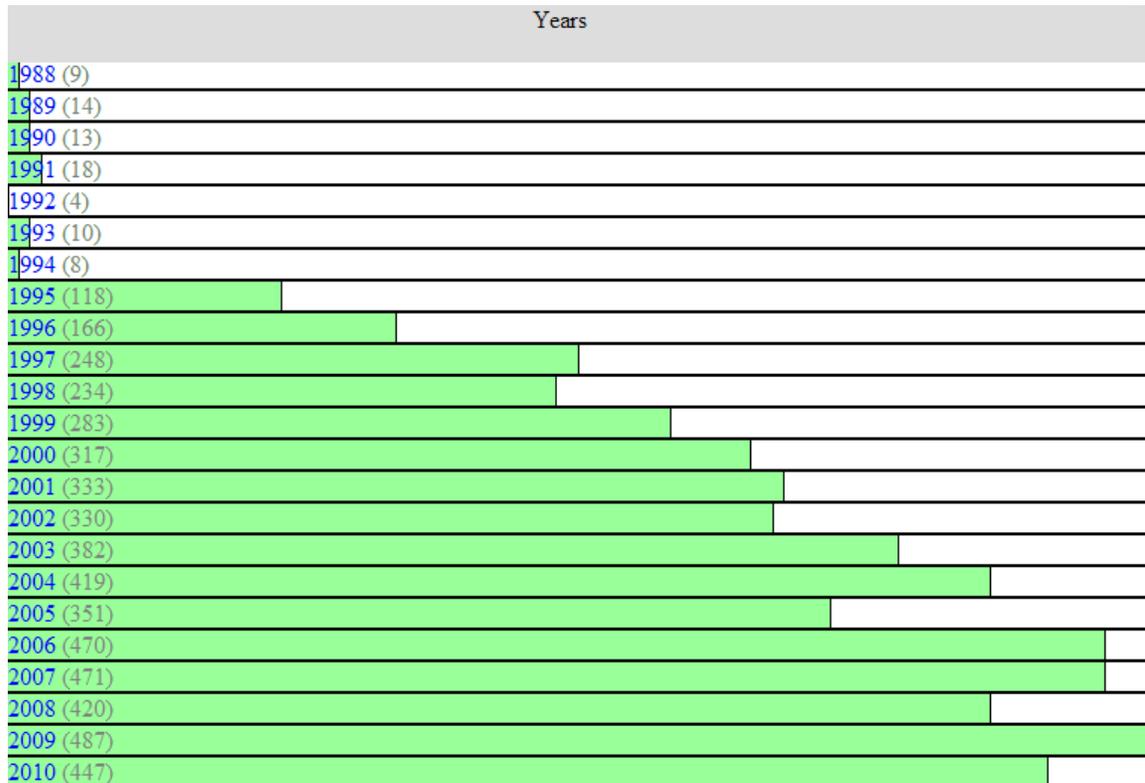

Figure 15.   Torpedo Ultra Search Engine results.

The chart pictured above allowed the user to look at subjects of interest by year. The chart displays years on the Y-axis and the length of the bars correspond to the embedded "hits." There is a dramatic jump of hits between 1994 (8) and 1995 (118). Had this search engine not been a U.S. Navy run application, these profound results would be interesting, but considering the source, it is probably safe to assume that the large jump is simply a result of better record keeping starting in 1995. However, the results from the Torpedo search engine serve a purpose. If the results of the journal-specific keyword search seemed underwhelming, this picture prompted a search-engine analysis that would perhaps be more convincing.

Since the slopes of the models for the journal-specific models are not particularly large (and sometimes not even statistically significant), it would seem as if this specific analysis fails to demonstrate much of a rise in the growth of simulation over time. However, Kelton's analysis of WSC attendance (Figure 14) and the search-engine wide results of Torpedo Ultra (Figure 15) suggest the possibility of a different means of displaying a rising trend. That trend becomes clear in Chapter III, which scours several scholarly search engines to show an impressive growth in simulation articles over time.

III.   SEARCH ENGINE KEYWORD ANALYSIS

I have every confidence that my simulation colleagues will continue to be able to use up whatever processing power is made available to them.

—Kelton (1994)

Being a computationally intensive technology, simulation has continued to take advantage of the improvements in processing speed over time.

—Hollocks (2006)

The results of the search engine data collection are the most striking and they truly display the impressive growth of simulation over the years. The "simulation" keyword results plotted against time is a convincing image, but this analysis goes further by fitting the data using the statistical software R. The purpose of this fit is not to predict the future, nor is it to prove the exponential nature of growth. Once again, the plot itself goes a long way in displaying the nature of the expansion. By fitting the data via regression in R, one can quantify the past growth and distinguish which years were particularly influential in the development and use of simulation as a research tool. These years that have been identified as influential will enrich the narrative and understanding of the sorts of advancements, both technological and theoretical, which have made the greatest impact on the simulation world. Since the data reflect when simulation-based articles have taken a jump or a decline in terms of publication, an influential year actually represents a window of time about which to investigate.

The remainder of this chapter is dedicated to the analysis of the data collected from the JSTOR, INFORMS, ACM DL, and EBSCOhost search engines. After a quick discussion of each search engine, the section displays the data and delves into analysis. The JSTOR analysis goes into great detail on the analysis methods that were used for each data set. As previously mentioned, analysis includes fitting a model and discussing influential years. Each analysis contains an Excel-generated scatter plot of the data that

compares the "simulation" data alongside "linear programming" and "optimization" data that were collected in the exact same matter as the simulation searches. The additional searches were to benchmark any trends in the simulation articles against other well-known operations research techniques. If the trends in the data were similar, for example, that would suggest that the growth in simulation was similar to the growth in OR as a whole, perhaps primarily reflecting advances in computing capability as well as general publishing trends within the academic community. It is important to note that "simulation" keyword searches were excluded from "optimization" keyword searches to avoid overlap in results that included articles with both keywords. This allowed for articles that dealt with optimization in simulation models to default as "simulation" articles.

A.   JSTOR

JSTOR is an online database and search engine whose access is granted through the Naval Postgraduate School. Its homepage claims that the database contains over 1,000 journals and its search engine appears as an option for any Operations Research related topic. The "Advanced Search" feature of JSTOR is simple and self-explanatory (see Figure 16). The ease of use created a sense of comfort in that the results would be accurate enough to trust the data.
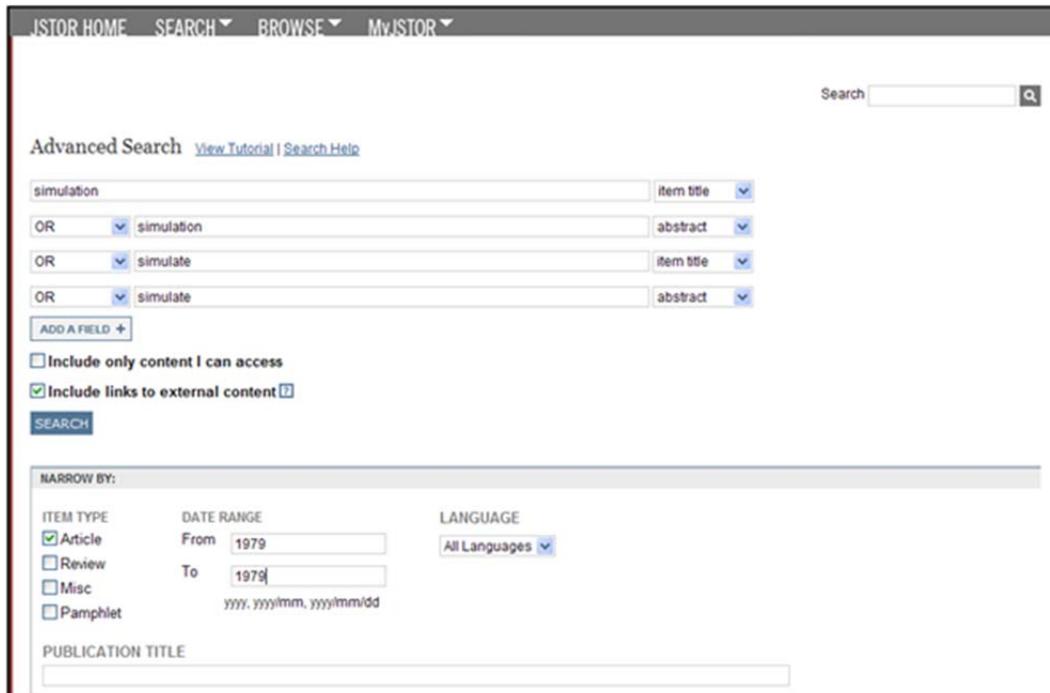
Figure 16.   Screenshot of the JSTOR "Advanced Search."

This search engine enables the user to see all articles that use specified terms from a variety of fields (such as "item title" or "abstract"). One also has the option to narrow results by year and by item type, which increases the accuracy of the results. By leaving the "Publication Title" field blank, appropriate articles are returned from all available journal

1.   Fitting the Model

The "Advanced Search" of the search engine was used to compile the data which resulted in a plot that, even without further analysis, reveals a tremendous positive trend in the number of articles in which "simulation" was a keyword (see Figure 17).
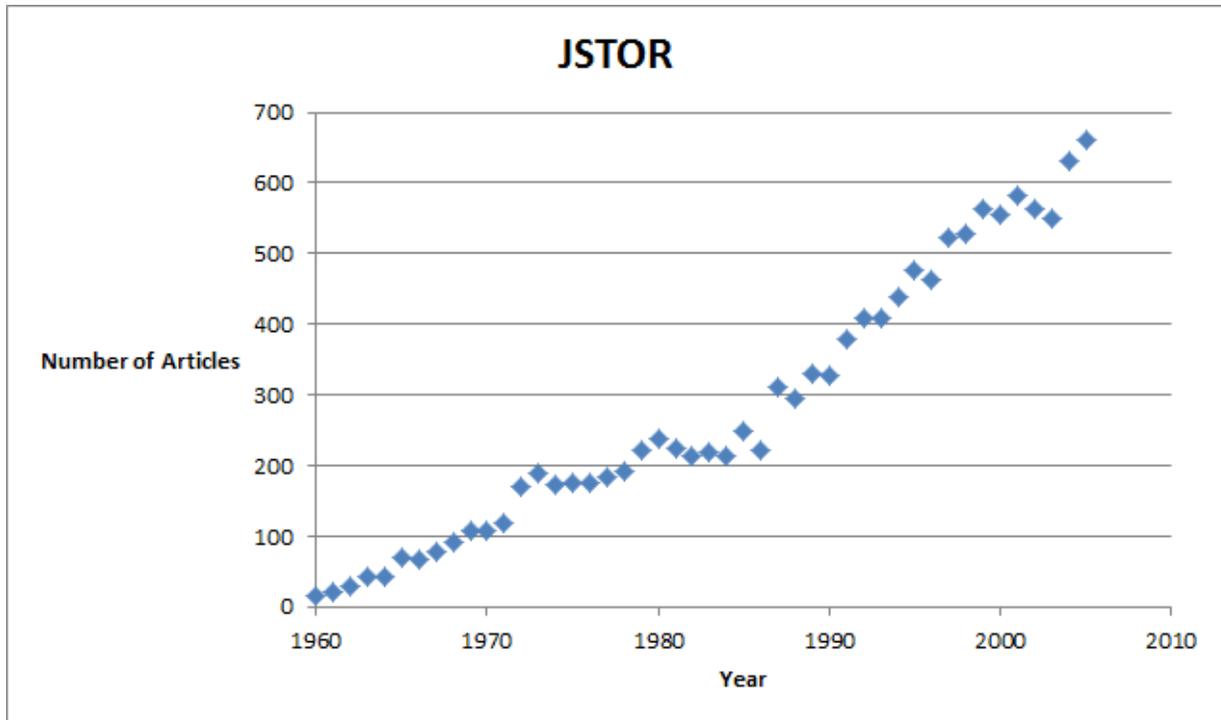
Figure 17.   Scatter plot of JSTOR data on simulation articles.

With this data set, as will be the case for all search engine data sets, analysis began with an initial, untransformed fit in R (see below).

- JSTOR_Sim.lm = lm(JSTOR_Sim~Year, data=engines.data)

From the above R entry, "JSTOR_Sim.lm" is simply a user-defined name for the model. The "lm" call is a function that creates a fit for the data with the response variable entered first (in this case, "JSTOR_Sim," the number of simulation based articles) and the independent variable(s) entered after the "~" (in this case, "Year," which is the year in question). "JSTOR_Sim" translates to "Frequency" in the final equations that will be presented in this section. The "engines.data" entry simply points to the data set of interest. It is important to note that 1960 is labeled "Year 1," and so on through 2010, which is "Year 51." The same formatting standard applies to all search engine regression fits. Once the model is in place, it is possible to view the summary statistics (see Figure 18).

42

```
Call:
lm(formula = JSTOR_Sim ~ Year, data = engines.data)

Residuals:
     Min       1Q    Median       3Q       Max
-104.262   -18.218    1.497   30.068    74.632

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -42.3623    12.1252  -3.494  0.00110 **
Year         13.6898     0.4492  30.474  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 40.45 on 44 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared: 0.9548      Adjusted R-squared: 0.9537
F-statistic: 928.6 on 1 and 44 DF,  p-value: < 2.2e-16
```

Figure 18.   R summary output of initial JSTOR data regression.

From the summary output above, the equation for the model can be derived:

- Frequency = -42.3623 + 13.6898 * Year

Without further analysis, it is obvious (thanks to the tiny p-value and the large R^2 value) that the frequency of the keyword "simulation" (and related words) is heavily dependent on time. The note "five observations deleted due to missingness" refers to the years 2006 through 2010, where the data were deleted due to clear unreliability in the JSTOR database. This unreliability is due to the fact that, at the time of this research, online availability was only complete through 2005. A look at the R residual plot confirmed the necessity of transformation in order to come to an appropriate fit for the data (see Figure 19).
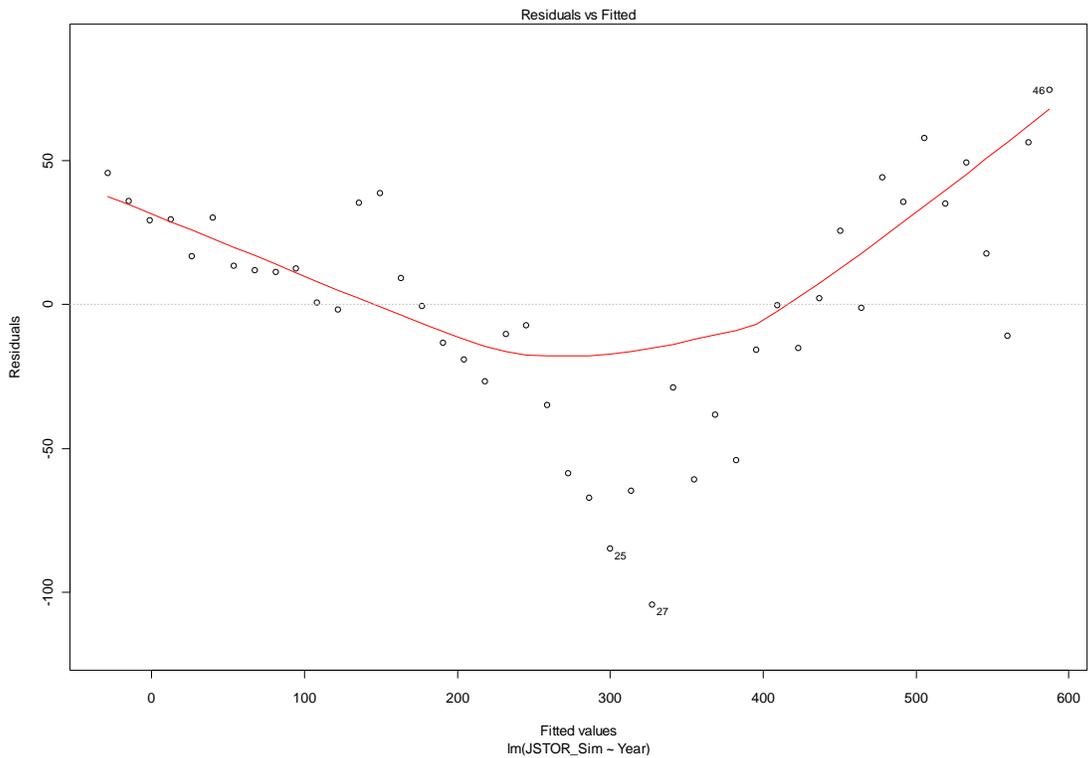
Figure 19.   R residuals plot of initial JSTOR data regression.

The heteroscedasticity in the above figure demands a transformation of the response variable. Notice once again that the plot displays some unusual behavior in the variance that remains unexplained. It has been suggested earlier in this thesis that editor interest or special "hot" topics may explain this phenomenon. Through trial and error, it was determined that the best transformation of the response variable was to take the square root of "JSTOR_Sim." This variable would appear in R as "JSTOR_Sim^.5". By regressing JSTOR_Sim$^{0.5}$ against Year, there is a dramatic improvement in heteroscedasticity (see Figure 20). Notice that two of the labeled data points that fall outside the otherwise equal variance are points 13 and 14. These data points will later be shown to be influential data points.
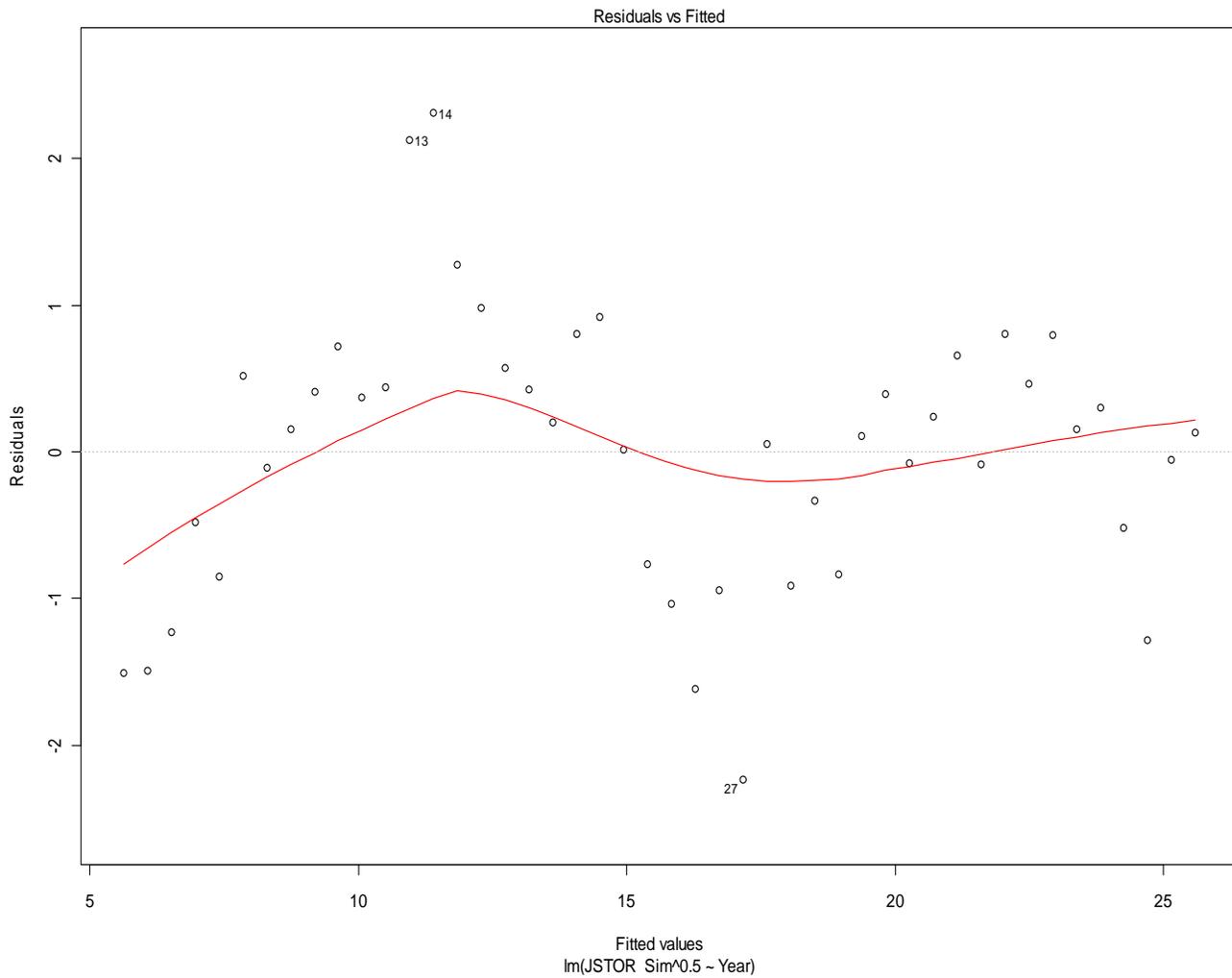
Figure 20.   R residuals plot of transformed JSTOR data regression.

The next step in the analysis is to look at the partial residual plot in order to determine if a transformation of the independent variable is necessary (see Figure 21). Specifically, is there a need for a polynomial term for "Year"?
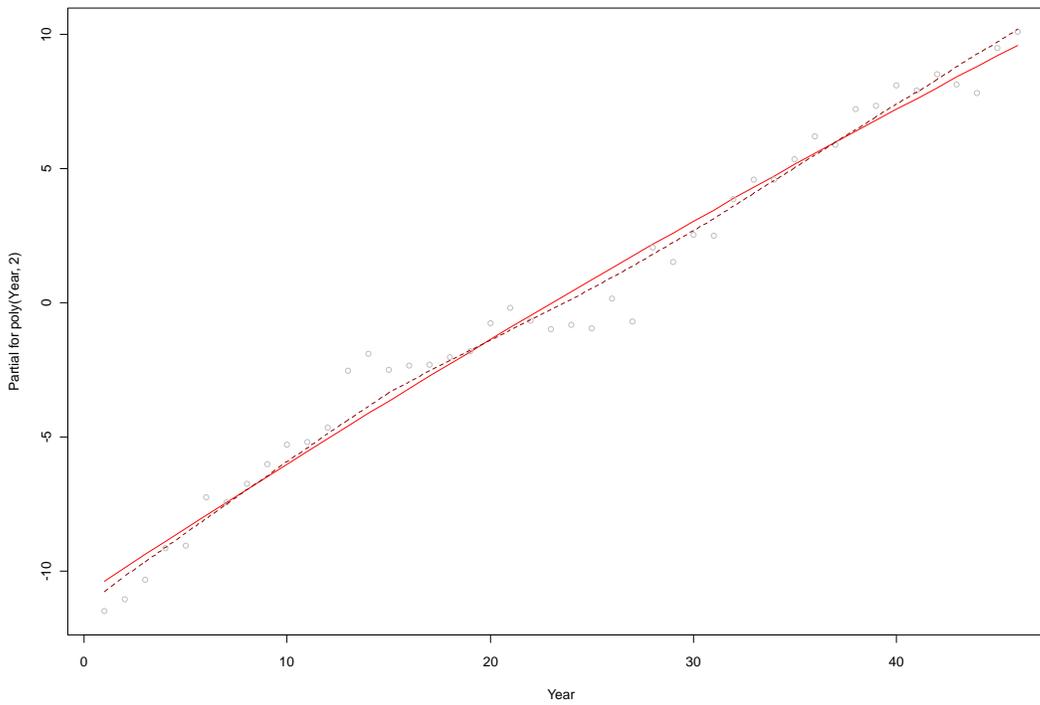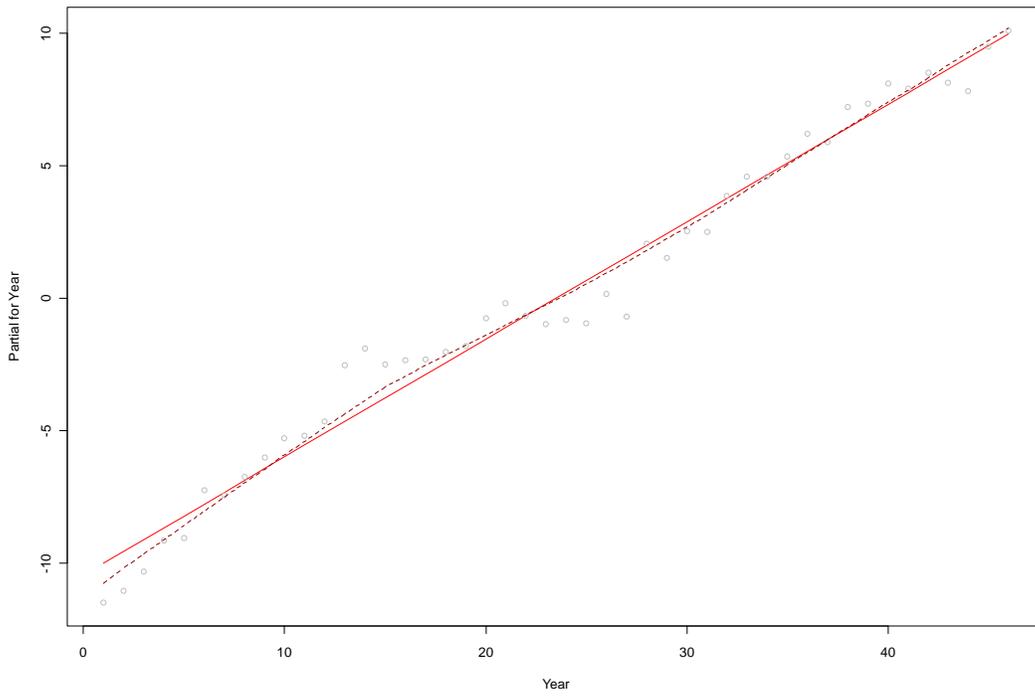
Figure 21.   R partial residual plots of JSTOR$^{0.5}$ ~ Year vs. JSTOR$^{0.5}$ ~ poly(Year,2).

The negligible difference between the above plots suggests no need to apply a polynomial to Year. Once again, the summary statistics of the model offer some insight into its validity (see Figure 22).

```
Call:
lm(formula = JSTOR_Sim^0.5 ~ Year, data = engines.data)

Residuals:
    Min      1Q  Median      3Q     Max
-2.2341 -0.7023  0.1415  0.5041  2.3123

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.18701    0.28125   18.44   <2e-16 ***
Year         0.44371    0.01042   42.58   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9382 on 44 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared: 0.9763      Adjusted R-squared: 0.9758
F-statistic:  1813 on 1 and 44 DF,  p-value: < 2.2e-16
```

Figure 22.   R summary output of JSTOR_Sim$^{0.5}$ ~ Year.

The final model for this data is as follows:

- Frequency $= (5.1870 + 0.4437 * \text{Year})^2$

The final model has a p-value < 0.0001, and an R^2 value of 0.976. Thus, nearly 98% of the variability is explained by this parsimonious model.

2.   Unusual and Influential Years

With a reasonable fit in place, it is possible to identify which data points (years) are particularly influential or unusual. This can be accomplished through use of R's Cook's Distance functionality (cooks.distance) and by comparing the results to a defined metric for what it means for a data point to be unusual (see Table 5). Cook's Distance

measures the impact of deleting a data point from a regression. Data points with a large Cook's distance merit closer examination (Cook, 1977).

Table 6.1 Approximate Cutoff Points for Identifyint "Unusual" Cases

Diagnostic Plots	Cutoff		
Partial regression plot	Visual inspection		
Partial residual plot	Visual inspection		
Univariate plots of			
h_i, e_i^*, DFITS$_i$, D$_i$, & DFBETAS$_{ij}$	Gaps and extreme values		

Diagnostic Statistics	Cutoff		
		low	*high*
Diagonal of hat matrix (h_i)		$2p/n$	$3p/n$
Studentized residuals ($|e_i^*|$)		$\alpha/2$	$\alpha/2n$
($|$RSTUDENT$|$)	t distribution (df $= n - p - 1$)		
$|$DFITS$_i|$		$2\sqrt{(p/n)}$	$\sqrt{p}$
Cook's D (D$_i$)		$4/n$	1
$|$DFBETAS$_{ij}|$		$2/\sqrt{n}$	1

NOTE: See text for discussion of cutoffs. For e_i^*, the "low" cutoff is for testing a single value of e_i^* and the "high" cutoff is for testing all n values of e_i^* with two-tailed tests.

Table 5.   Numerous ways in which to determine unusual cases (From Bollen and Jackman, 1990).

The "cutoff" for unusual data points for Cook's D_i is 4/n, where "n" is the number of data points. In the case of the JSTOR database: $4/n = 4/46 = 0.0869$. A couple of lines in R can determine which points are beyond the cutoff for unusual data. In the JSTOR_Sim model, the "unusual" Years are 1, 2, 13 and 14 (1960, 1961, 1972 and 1973).

The years 1960 and 1961 represent a time that was discussed in Chapter 1 of this thesis and a further discussion of those years will appear in the EBSCOhost analysis. The remaining unusual years are 1972 and 1973.   Upon inspection of the data, one can see that there was a particularly high jump in simulation articles between 1971 and 1972 (120 articles in 1971 and 171 articles in 1972). One would expect to find some interesting happenings that would pertain to simulation on or about the year 1971. A simple internet search reveals a few possibilities (from www.computerhope.com):

48

- 1970: Intel announced the first RAM chip, the Intel 1103, with more than 1,000 bits of memory.

- 1971: Intel developed the first microprocessor, the Intel 4004, capable of 60,000 instructions per second and a clock speed of 740 kHz.

- 1972: The C programming language was developed. It introduced structured programming and was a prelude to object-oriented programming.

Further scholarly research suggests some other landmarks that very well may have contributed to a surge in simulation use.

- 1969: Alan Pritsker of Purdue University produced GASP II, software that aided in making portable simulation programming languages (Hollocks, 2006). At the time, this FORTRAN-based language had been praised for being flexible, well documented and, while it was relatively new and unknown, it was already being taught at Arizona State University (Petersen, 1969).

- 1970: B. W. Hollocks implemented GSP-III's features with FORTRAN as a platform to create FORSS (FORTRAN-based Simulation System). Its portability resulted in wide use (Hollocks, 2006).

- 1970: Jeffrey R. Raskin's "A Tutorial on Random Number Generation" describes the benefits of computer-generated pseudorandom numbers, particularly their ease of generation and the ability to reproduce results (Raskin, 1970). Raskin's paper represents one of many at the time that dealt with the idea of improving on existing methods of random number generation because of its necessity in the field of simulation.

3.   Conclusions

The analysis of the JSTOR data identified an influential window of time from the late 1960s through the early 1970s. Computational developments have almost certainly played a role in the apparent jump of simulation-inspired articles. The growth of simulation articles over time reveals a convincing trend, and the plots of "optimization" and "linear programming" keyword searches are available for comparison (see Figure 23).
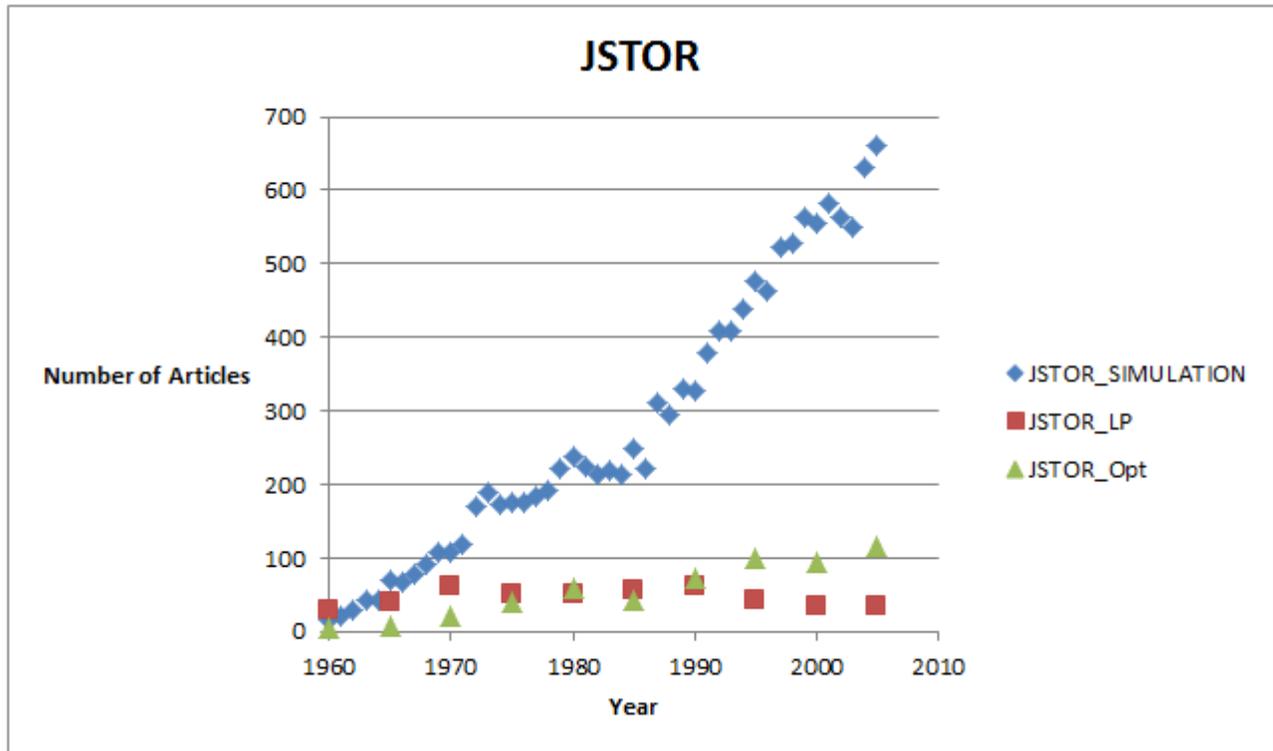
Figure 23. JSTOR data with "optimization" and "linear programming" results.

The growth of the simulation data is impressive when compared to the slight decline of linear programming (LP), and the slow linear growth of optimization (Opt) articles. For the JSTOR search engine, it was possible to exclude results that contained "simulation" as a keyword in the "optimization" search, which is necessary to ensure that articles that dealt with optimization in simulation are categorized correctly. Overall, the growth during the first 30 years is less than half that of the last 20 years. The changes in JSTOR_LP and JSTOR_Opt are not nearly so dramatic. JSTOR_LP rises slightly but then falls near its original levels, while JSTOR_Opt increases at a slow but steady rate.

The analysis of the JSTOR search engine was presented in full as a means of demonstrating how analysis was applied to all search engines. The data were fit to a regression, and from that regression, significant data points were identified. The results and conclusions of the remaining databases will not be presented step-by-step. Rather, the final model and scatter plot will be revealed and the influential years discussed.

B.   INFORMS

The INFORMS search engine is available on the INFORMS Online website. It contains a complete database of articles dating from before 1960 and allows access to all articles published after 1998. It is simple and user-friendly, but the results are limited to those articles that were published within INFORMS (see Figure 24).

Figure 24.   Screenshot of the INFORMS "Advanced Search"

While the process of using the search engine was simple enough to create a trustworthy database, the fact that the results only reflect INFORMS publications produces a less striking growth than the other search engines

1. Fitting the Model

With data collected for all years between 1960 and 2010 in the INFORMS data set it was possible to recognize particularly large jumps (or drops) in simulation article frequency. This also allowed for enough data to be collected to demonstrate a particular trend as well as enough data to allow for a fit that would reveal influential years. Like most data sets, specific inference cannot be derived from simply looking at the plot. While a dramatic increase in annual articles suggests a growth that is quicker than linear, such a claim is not obvious in the INFORMS data (see Figure 25).
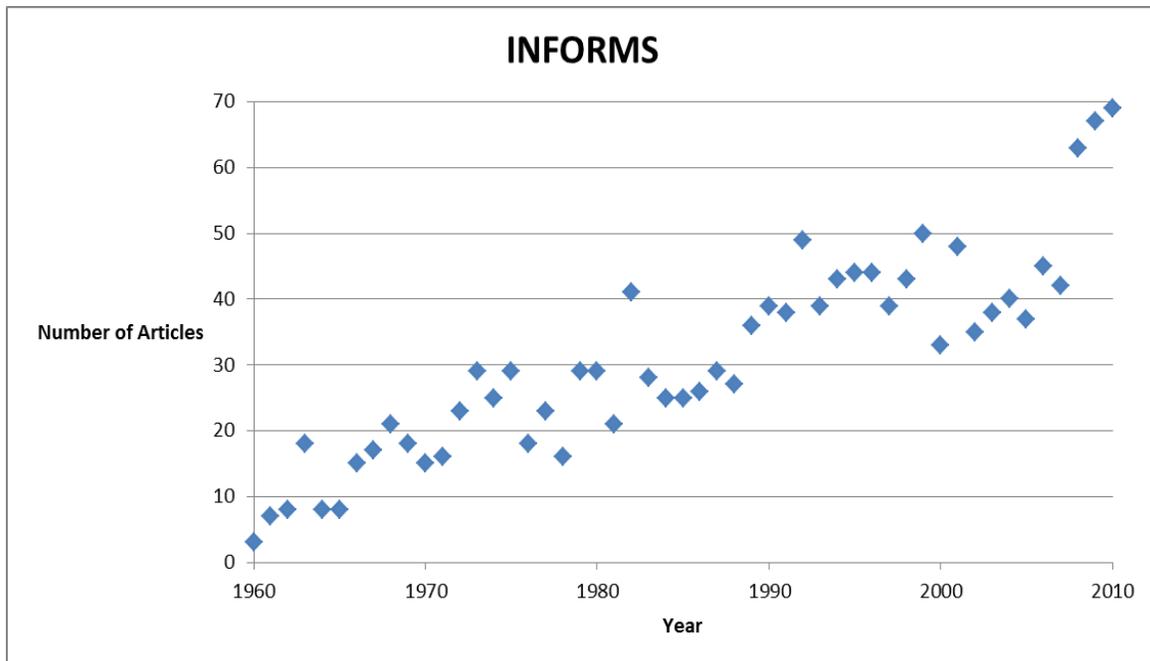

Figure 25.   Scatter plot of INFORMS data on simulation articles.

After following a procedure similar to that of the JSTOR data, the final model for the INFORMS data is INFORMS_Sim$^{0.4}$ ~ poly(Year, 2), which translates as follows:

- Frequency$^{0.4}$ = 3.8179 + (5.2906 * Year) - (0.8676 * Year^2)

This model has a p-value < 0.0004 for the first-order term for Year, a p-value < 0.05 for the second-order term and an R^2 value of 0.822.

2.   Unusual and Influential Years

The same "Cook's Distance" method that was used to determine influential data points in the previous analysis was used for the INFORMS model. The influential years were 1960, 1963, and 2008 through 2010. This could simply because these are a set of "bookend" years, or because they were truly unusual. By the year 2008, advances in computing technology are so dense that it is nearly impossible to pinpoint several events that could have inspired a leap. However, a look at the data as represented on a scatterplot reveal an interesting story. That scatterplot will be discussed in the "conclusions" section.

3.   Conclusions

While the numerical results of the INFORMS search engine may appear to be less impressive, the fact that the search results spawn from a smaller sample of publications allows for a more intimate analysis. Also, the smaller values inherent to the data set do not stop an obvious difference in the trends of simulation articles and linear programming articles. Optimization and linear programming data were not collected for all years between 1960 and 2010 in any of the data sets. This data were collected in an iterative process that called for collection in "clumps" of data if the model showed a window of unusual observations (see Figure 26).
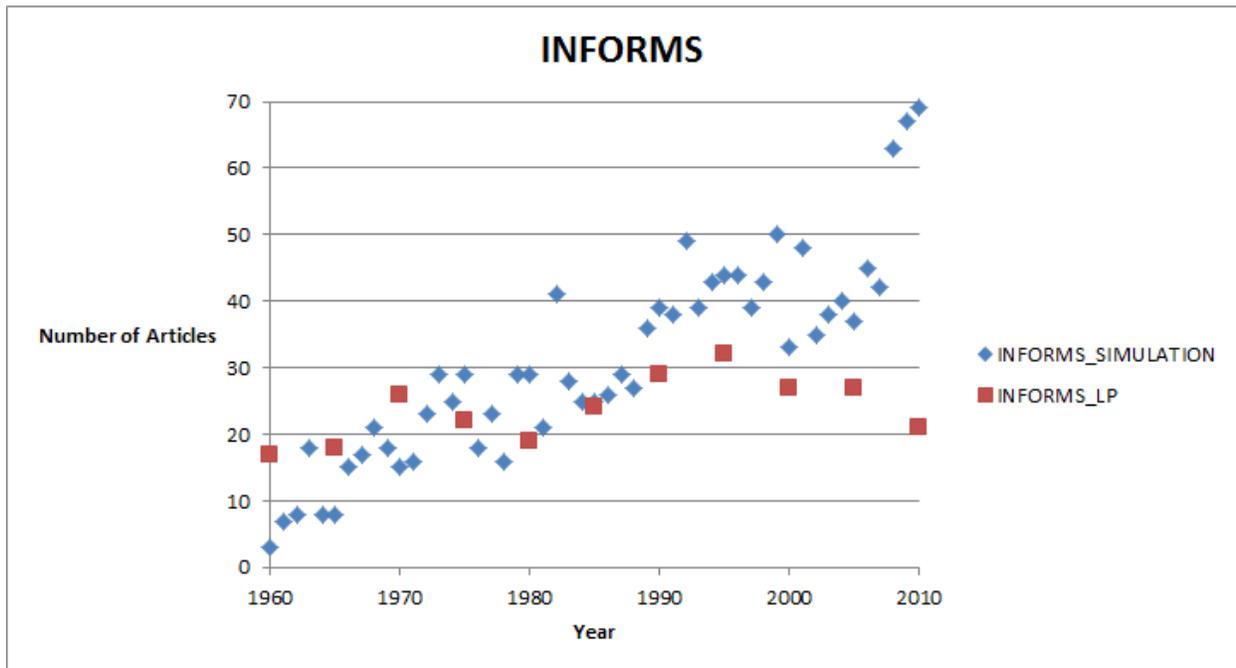
Figure 26.   INFORMS data with "linear programming" results.

Once again, the simulation results grow faster than linearly while the LP results display a linear trend. Hypothesis testing confirms the linearity of the "LP" data. The search engine did not allow for exclusion of "simulation" from "optimization," so no trend line for "optimization" exists. One of the most interesting aspects of the simulation curve is how the data points drop below the apparent trend line for the years 2002 – 2007. An exploration into the INFORMS society itself may explain the drop. The following are summaries of key events obtained from the INFORMS Simulation Society Business Meetings minutes from 1999 through 2008.

- May 1999: At this time, the organization was known as the INFORMS College on Simulation. The College was acting like a Society in that, among other things, it was running the PhD colloquium at the Winter Simulation Conference and actively inviting interested students to attend the conference. The minutes reveal a decline in College representation at the annual meetings and a need to encourage more simulation related article submissions.

- Nov. 1999: More concern over the dropping number of simulation submissions was noted. It is mentioned that the College would have more influence if it were a society.

- Oct. 2004: The College was over 500 members, which was large enough to be considered a Society. The members at the meeting unanimously vote on a proposal to petition for Society status.

- Nov. 2005: The organization held its first meeting as a Society. It is mentioned that simulation submissions are on the rise.

- Nov. 2007. As a sign of great success, it is mentioned that 742 authors are contributing to the Winter Simulation Conference, and that all the hotel blocks reserved for the conference have been booked.

By 2008, the number of simulation related articles have risen above the trend line. Perhaps it is no coincidence that the increasing trend in simulation articles found in Informs journals matches the rise of "Impact Factors" for the *INFORMS Journal on Computing* (see Ch. 2).

C.   ACM DL

The ACM DL (Digital Library) website search engine offers access to every article ever published by ACM. Searching for individual articles by author or title is simple, and the keyword search is straightforward. The advanced search option can be a bit confusing, but for the purposes of a keyword search for "simulation," confusion can be avoided by refining a simple keyword search by publication date. The database was created with this simple method. Furthermore, the data set was intended to be small (therefore incomplete), as this particular analysis was meant to show a consistency in the rapid growth of simulation use across search engines. Figure 27 represents the results for a "simulation" keyword search for all publications in the year 1970.

Figure 27. Screenshot of the ACM DL keyword search.

The complicated "Advanced Search" option may have resulted in a more comprehensive database, but this simple search method managed to create a very compelling outcome, one that will be discussed in the next section.

1. Fitting the Model

The data set for the ACM DL search engine hints at an exponential growth in simulation articles before detailed analysis is applied (see Figure 28).
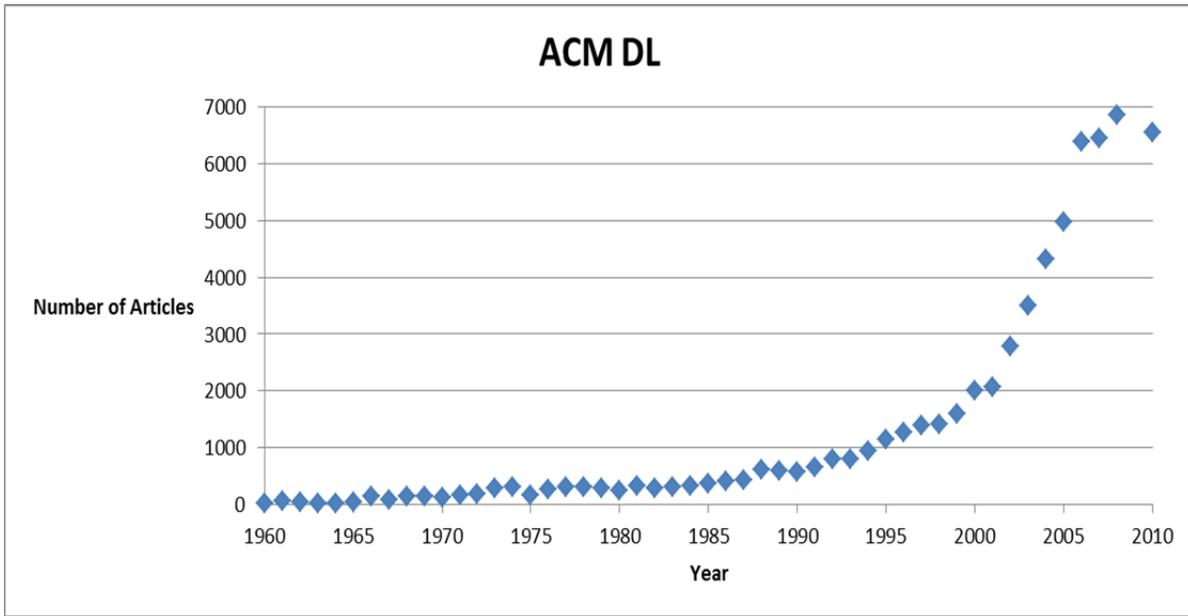
Figure 28.   Scatter plot of ACM DL data on simulation articles.

While the earlier results show polynomial growth in the number of simulation-related articles, the results for the ACM DL search engine data are fit well by an exponential curve (p-value < 0.0001, $R^2 = 0.945$). The final model is a first-order model for the response log(ACM_DL_Sim), which translates as follows:

- log(Frequency) = 3.4251 + (0.1051 * Year).

2.   Unusual and Influential Years

The influential years (in terms of Cook's Distance) are 1960, 1964, 1966, and 1967. This set of years shall be discussed in the results of the next search engine, but a look at the scatter plot of log(Frequency) clearly displays unusual behavior in the early years of the data (see Figure 29).
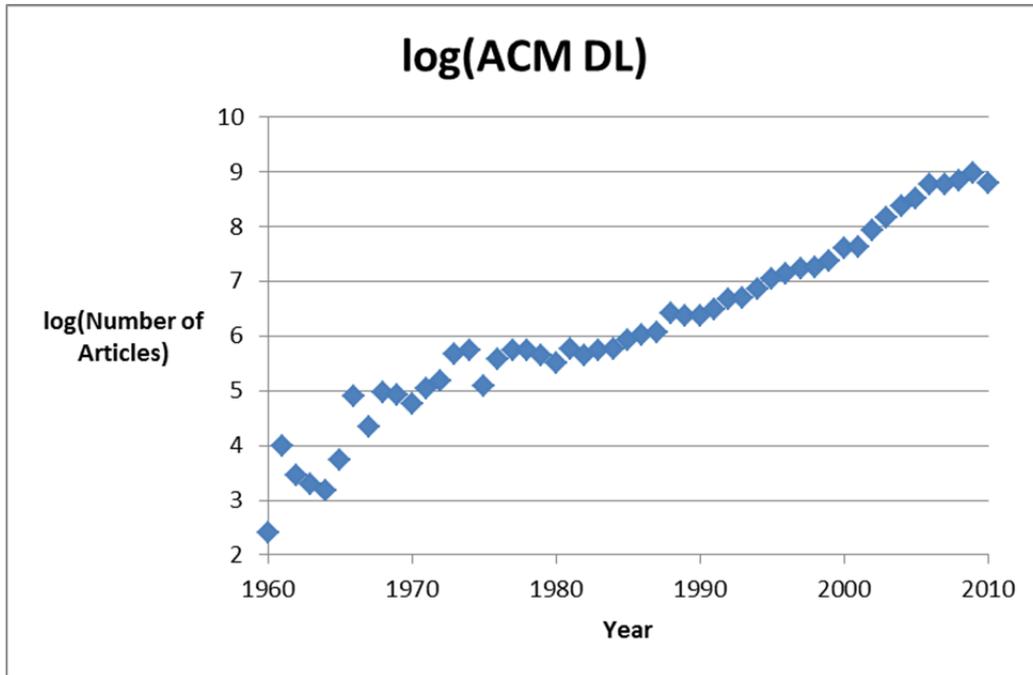
Figure 29.   Scatter plot of log(Frequency) of ACM DLdata on simulation articles.

3.   Conclusions

The ACM DL search engine allowed for separation of "simulation" from "optimization" searches.   The "optimization" (Opt) and "linear programming" (LP) curves, while exponential (as confirmed by identical analysis techniques), are dwarfed by the "simulation" (Sim) curve (see Figure 30).
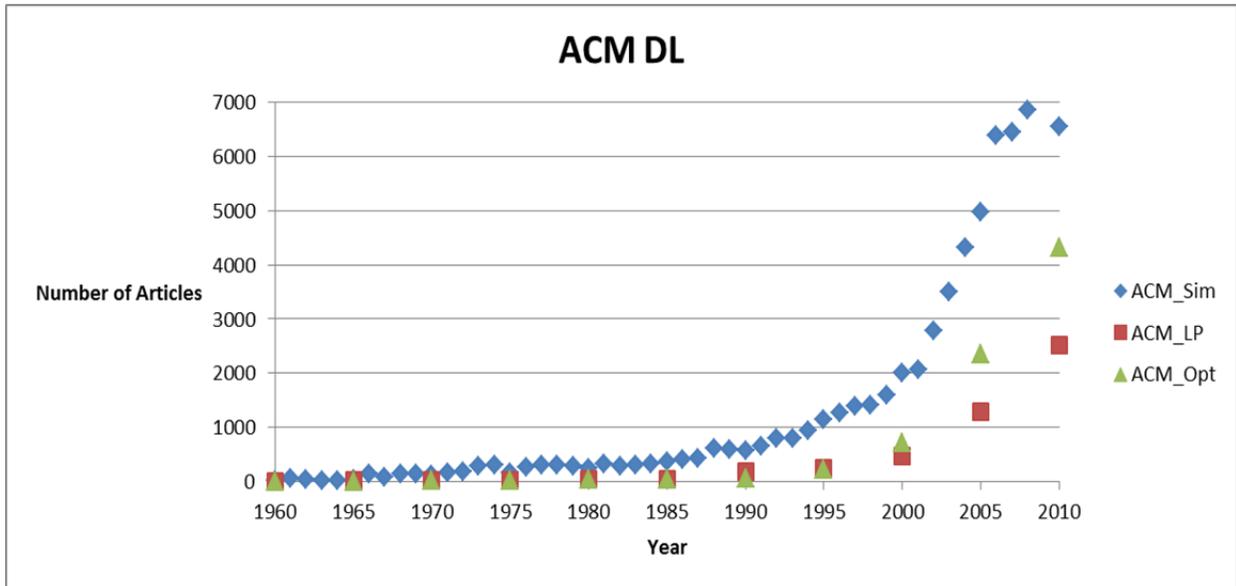
Figure 30. ACM DL data with "optimization" and "linear programming" results.

A look at the exponential curve and the data point corresponding to the year 2005 reveals some interesting characteristics. The change from 2000–2005 is larger than that from 2005–2010, while the 2010 data point falls below the curve. This might suggest that not much has happened between 2005 and 2010 as far as computing capability is concerned, and the world may be poised for another advancement! The rather large leap in the plot between 2000 and 2005 inspires an investigation of computing history as far back as 2000.

- 2000: Intel releases the Pentium 4.
- 2003: Intel releases Pentium M.
- 2005: Intel releases Pentium D.
- 2006: Intel releases Core 2.
- 2007–2009: Intel releases several advancements to the Core2.

These advancements, while rapid and significant, do not relate specifically to the world of simulation as to suggest that they play the most important role in simulation's growth. In fact, the years 2005 and 2010 have not been identified as influential. The advancements listed have certainly improved computers as far as processing power and overall speed are concerned, but perhaps the simulation community is about to witness a breakthrough that is more than a leap in technology.

D.   EBSCOHOST

The EBSCOhost online databases provide access to publications for various fields of research. Any "operations research" search executed from the NPS Library website leads to the EBSCOhost "Business Source Complete" database. This database offers references to over 1,300 scholarly journals. The advanced search engine is the easiest and most self-explanatory, and the articles are updated frequently (see Figure 31). The ease of use inspired a comprehensive study in the later years whose results are very convincing.

Figure 31.   Screenshot of the EBSCOhost Business Source Complete keyword search.

This search engine could exclude search terms so that "optimization" and "simulation" results would not be mixed. Furthermore, the ability to "apply related keywords" allows the user to only search with a single word instead of deciding which other words may apply.

1.   Fitting the Model

This data set displays a linear-looking (yet uphill) trend from the years 1960–1985. However, it is easy to see that things get very interesting after the year 1990 (see Figure 32).
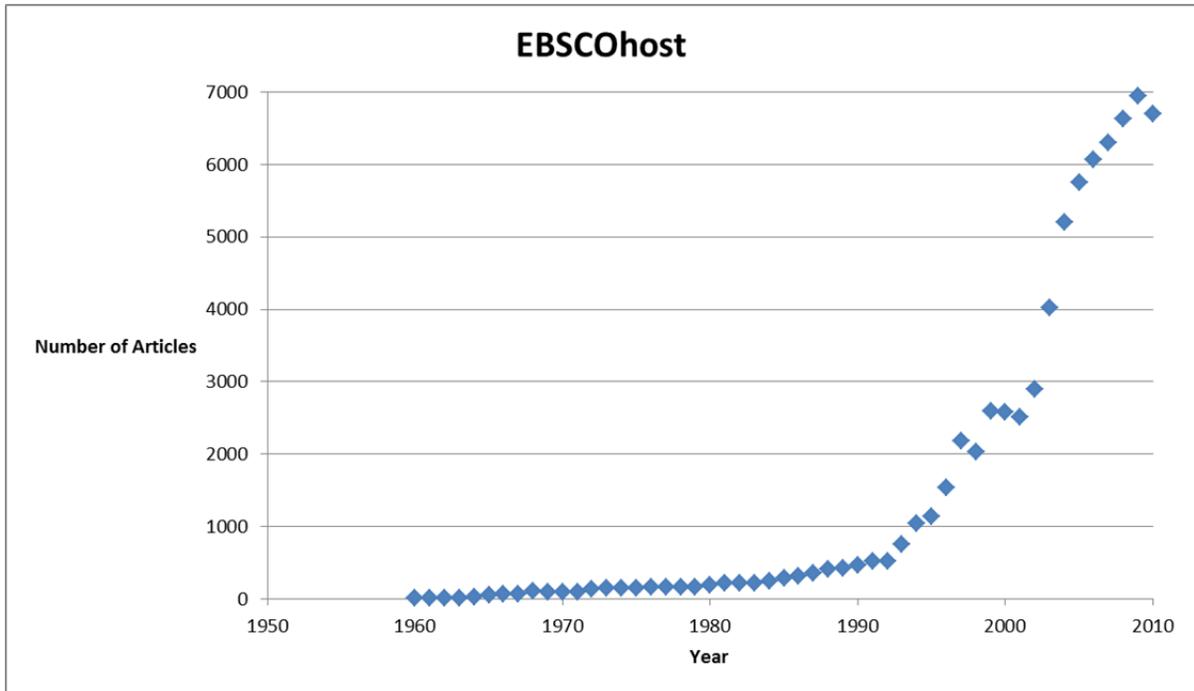
Figure 32.   Scatter plot of EBSCOhost data on simulation articles.

A linear model with log(EBSCOhost_Sim) as the response fits well (p-value < 0.0001, R^2 = 0.966). The model translates as follows:

- log(Frequency) = 2.8499 + 0.1194 * Year

2.   Unusual and Influential Years

The influential years (according to Cook's distance) are 1961, 1962, 1966 and 1968; 1963 (with Cook's distance 0.0768) is also very close to the influential cutoff ($4/n$ = 0.0784). This set of years is remarkable in that they represent early days of computer simulation other than the "bookend" year of 1960. This particular analysis motivates a look at developments from the early to late 1960s that would have been influential, and those developments are discussed in detail in the "conclusions" section.

3. Conclusions

The EBSCOhost search engine allowed for separation of "simulation" from "optimization" searches, which allowed for a comparison of simulation, optimization, and linear programming data on the same plot (see Figure 33).
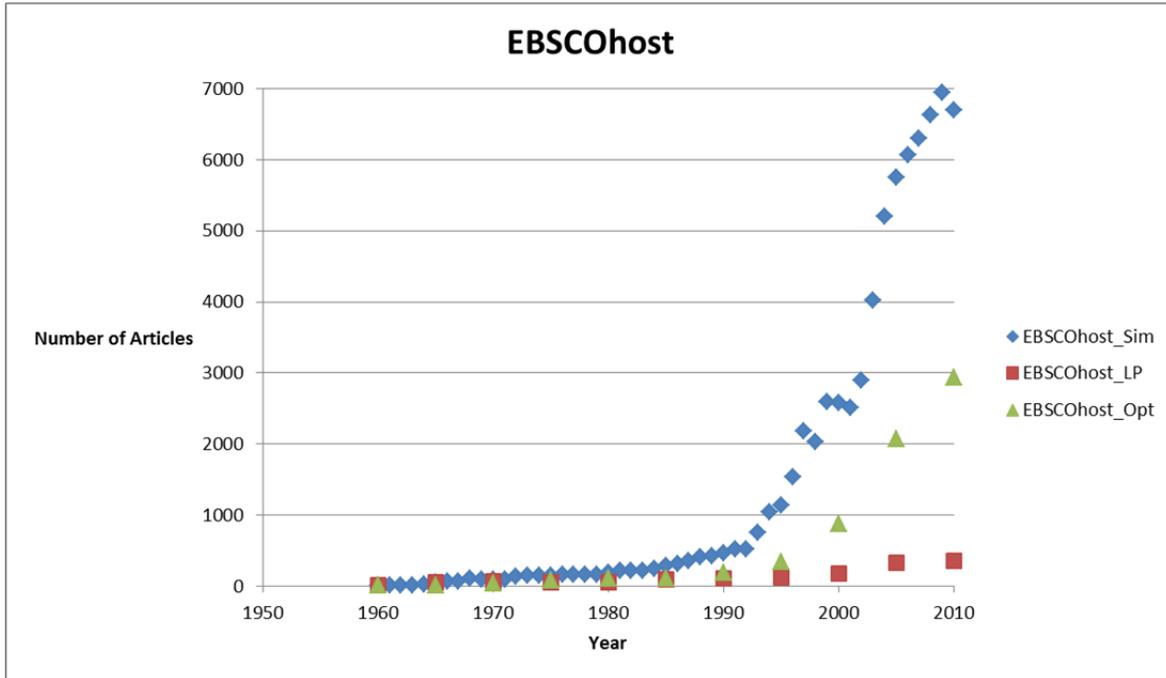
Figure 33.   EBSCOhost data with "optimization" and "linear programming" results.

The "optimization" (Opt) curve, while representing smaller return values, also fits an exponential model.   Hypothesis testing confirms the linearity of the "linear programming" (LP) data.

With 1961 through 1963, 1966, and 1968 identified as the influential years for this fit, it was necessary to look at developments in the computing world *and* the world of academia as it pertained to simulation. This necessity is because, with hindsight being what it is, common knowledge tells us that technological advances could not have been the only driving force behind the advancement of simulation modeling in the 1960s.

- 1961: The programming language FORTRAN IV is created.
- 1963: IEEE is founded. IEEE will be responsible for publishing many simulation based articles.

- 1964: BASIC is run for the first time.
- 1964: IBM introduces System/360, which uses interchangeable software. This is an example of a "third generation computer."
- 1965: Gordon Moore publishes Moore's law.
- 1969: Peter A. W. Lewis publishes his paper, "A pseudo-random number generator for the System/360." This paper describes a "pseudo-random number generator that uses the full capacity of the 32-bit registers of IBM SYSTEM/360 computers," thereby addressing a hot topic of the day that was common to critiques of simulation (Lewis, Goodman & Miller, 1969).

The efforts of Peter A. W. Lewis were not the first attempt to develop random numbers and variates that behaved according to accepted statistical properties. Such efforts dated as far back as 1949 with "the multiplicative congruential generator originally introduced by Lehmer." In fact, by 1971, over 200 papers had been written on the subject of random/pseudorandom number generation (Craddock & Farmer, 1971). This amount of publications was indicative of the magnitude of attention that stochastic modeling was receiving. So, while he may not have broken ground in the random number generation field, Dr. Lewis was a pioneer in the simulation community mostly because he was a respected scholar who took it seriously.

As discussed in Chapter 1 of this thesis, one of the most prohibitive factors that had been holding back the advancement of simulation was the negative attitude that the Operational Research community expressed towards it. Dr. Lewis was one of the first highly respected scholars to visualize simulation as a valid method for modeling and research. His endorsement played a tremendous role in the eventual acceptance of simulation as a noble pursuit. Dr. Lewis's impact was so great in fact that he was recently awarded the 2012 INFORMS Lifetime Professional Achievement Award (LPAA) in recognition of his contributions. His nomination letter, graciously provided by Professor Lee W. Schruben of the University of California, Berkeley, perfectly expresses the way in which Dr. Lewis lit the fuse on what would become an explosive discipline in the world of Operations Research.

Lee W. Schruben, Chancellor's Professor
Department Of Industrial Engineering
   And Operations Research
College Of Engineering
4117 Etcheverry Hall
Berkeley, California 94720

Phone
Fax

MEMO TO: I-SIM Lifetime Professional Achievement Award Committee
                    c/o Dr. Averill M. Law: Chair

FROM: Lee Schruben
                   Professor, Industrial Engineering & Operations Research and Informs Fellow

DATE: August 17, 2011

SUBJECT: Nomination of Peter A. W. Lewis for 2012 LPAA

I wish to nominate Professor Peter A. W. Lewis for the INFORMS Simulation Society Lifetime Professional Achievement Award, the highest honor given by the Society. Professor Lewis unfortunately died on April 8th of this year, after this nomination process had been started. Those of us involved unanimously decided after his death to proceed with this nomination. However, this did not begin, nor is it now a "sympathy" nomination; it is simply the right thing to do, sadly in my opinion coming late, but still right. Fortunately, Peter had already given me permission to nominate him so he knew about it. I was told by his son, Vlad, that Peter was very pleased to be remembered so long after his retirement and lengthy illness, which he suffered mostly in seclusion.

It is the belief of those of us who initiated this nomination that it is important for future generations in our field to understand that simulation was not always considered a respected, or even welcome, academic discipline in Operations Research. Indeed, it now seems shocking that at least two previous LPAA award winners, our most famous people, were actually denied tenure in their first academic appointments – not because of the importance and quality of their work, but because of *what* they did – simulation research. Among their senior colleagues simulation research had not yet been established as a credible intellectual discipline in Operations Research. There were simply <u>no</u> famous academics in the world willing to write supporting tenure letters for Assistant Professors of Simulation.

It was only until scholars, Peter Lewis prominently among them, who were already world-famous academics in other fields, embraced simulation as an intellectual research field was it even possible for academics like myself to name anyone who our colleagues respected to write a tenure support letter. Peter Lewis had recently expanded his research portfolio to include "simulation". This was in spite of simulation being considered pedestrian by all of his more-puritan mathematical colleagues in operations research. I was one of the first Assistant Professors at in a top-tier Operations Research department (Cornell) to receive tenure in my first job because academic giants like Peter Lewis, who were highly respected by my colleagues, were willing to take simulation research seriously and write a supporting tenure letter. (A year earlier would have been too soon!) I and others in my generation in Operations Research write many successful tenure support letters every year, and we would not be in a position to do this had not Peter Lewis, an already famous scholar, first done so for me and one or two others my age. I was informed at my mid-tenure review, and by all my senior mentors,

that I would not have any chance for tenure in Operations Research if I continued doing "simulation methodology research". Simulation was viewed as mere computer programs used as the method the last resort by the mathematically inept. "The Method of Last Resort" is literally the title of the Simulation Chapter in the most-used Operations Research textbook of the day (that by Harvey Wagner). Peter Lewis alone had the personal stature to persuade all the statisticians at Cornell that simulation should indeed to be taken seriously and supported my tenure case. I was also lucky that Cornell did not have a separate Statistics Department at the time and statisticians were in, and dominated, the School of Operations Research and Industrial Engineering.

I firmly believe that future generations in our field should know that simulation research would still not exist as an academic field in Operations Research today without Peter Lewis, who, facing derision of some of his own Operations Research colleagues at NPS, recognized simulation research as having scholarly value.

Simulation accomplishments by Peter Lewis include early, and still in use, random number generation methods, random variable generation methods, and laying mathematical foundations for generating multivariate time series and analysis of renewal processes. He wrote several deep textbooks on Simulation and Monte Carlo Methods… (this section to be somewhat fleshed out later).

At a critical time, Peter Lewis played a central role in winning the support of the American Statistical Association to co-sponsor the Winter Simulation Conference, and served as its representative on the WSC Board of Directors.

Several supporting letters are included from across a broad spectrum of leaders in Simulation Operations Research. Since I know you will read them carefully, I will not repeat them here. Only emphasize that each of these letters, in my opinion, is sufficient to justify selecting Prof. Lewis for this honor; he exceeded all criteria. This is historically important: we need to remember those early giants on whose shoulders our field now stands.

It is perhaps not too much to claim that Operations Research academics in simulation analysis methodology and everybody in our field who uses simulation statistical methodology owe their careers in some degree to Peter A. W. Lewis. He provided a critical link, a tipping point, in the intellectual evolution of Simulation in Operations Research. That is why the highest INFORMS Simulation Society award is so appropriate for Peter A. W. Lewis. Operations Research might not yet even have a Simulation Society without him.

Attachment 1: INFORMS LPAA Nomination Letter on behalf of Peter A. W. Lewis.

Prof. Schruben's nomination letter tells the story of a time when simulation had yet to gain respect in the OR community. Dr. Lewis's endorsement of the subject allowed for an acceleration in scholarly pursuits of advancing computer simulation. In order for the once-controversial subject of simulation to soar to greater heights, it first had to get off the ground. Dr. Lewis represented one man in a relatively small class of respected scholars that allowed for that to happen.

THIS PAGE INTENTIONALLY LEFT BLANK

IV.   CONCLUSIONS

Not that evolution is over for simulation; there's life in the old technique yet—and that may include surprises.

—Hollocks (2006)

A.   SUMMARY OF EXPONENTIAL SIMULATION GROWTH CURVES

The previous chapter discussed the data that were collected on four scholarly search engines and the models that resulted from regression analysis. The four data sets resulted in four separate and distinct models that were used to prove the rapidly increasing trends of simulation use as well as determining the most influential years in simulation history. While some of the models share influential years, they are otherwise different models. Nonetheless, all show a greater-than-linear or exponential growth. This can be seen with a review of each of the plots from the four search-engine based data sets (see Figure 34).
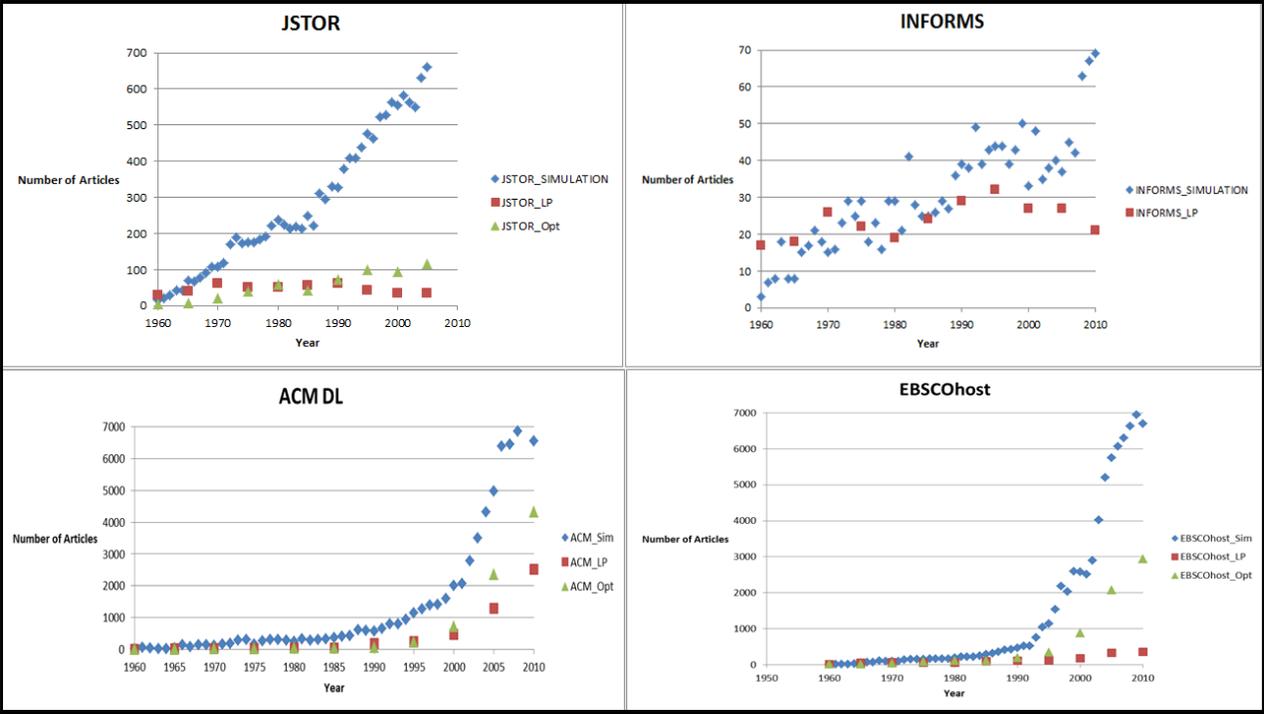
Figure 34.   Search Engine scatter plots for each model.

The rapid growth demonstrated by each model contributes some validity to a theory offered by B. W. Hollocks in his 2006 paper, "Forty years of discrete-event simulation—a personal reflection." Hollocks provided an in-depth look at technological advances that he has observed throughout his time in the OR community. He suggested a tweak to Moore's Law as it applies to simulation. Instead of a trend based on values that would double every 18 months "until constrained by the laws of physics," Hollocks predicted that simulation's evolution would follow more of an "S-curve" (see Figure 35) (Hollocks, 2006).
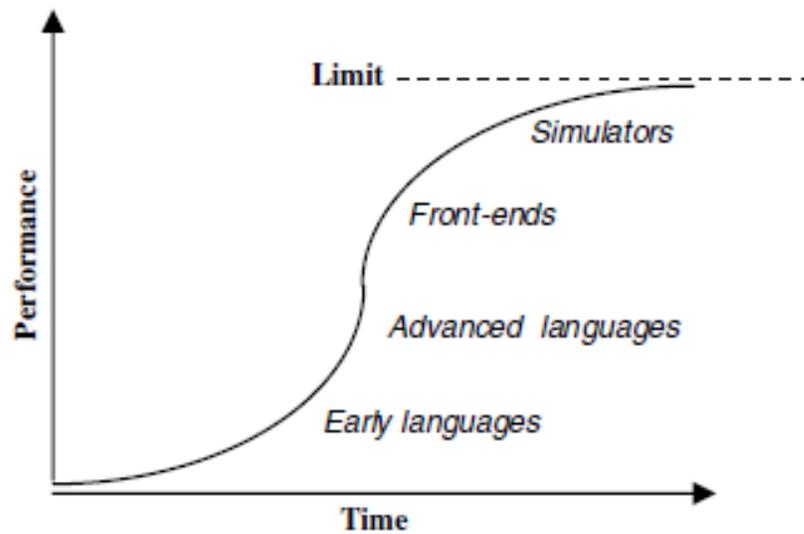
Figure 35.   S-curve of simulation growth (From Hollocks, 2006).

The curve proposed by Hollocks says that as computing capability reaches its physical upper-bound, a bound that Gordon Moore himself predicted would occur around 2015 to 2020, limited scope for improvement might be left (Saran, 2005). What room does remain may be user-oriented improvements such as ease of model building or display capabilities. Such a theory, however, would mean that simulation has enjoyed success and expansion over time solely because of technological advances. A "squinty-eyed" look at the curves presented in this research may suggest that Hollocks is absolutely correct in giving so much credit to improvements in processing speed and software. Of advanced processing speed, Hollocks said that "in addition to permitting the same tools to run faster, it has also permitted the tools to carry more function and feature." Simply put, Hollocks is suggesting that the simulation field is getting very close to being as advanced as it will ever be (Hollocks, 2006). The author of *this* research has another suggestion, which seems to agree with Kelton (1994) when he all but predicted the future of simulation advances. Kelton suggested, and this research shows, that advances in simulation have been both technological *and* theoretical. He said that

"experimental design methods" were "ripe for accelerated research progress," but that for the impact of simulation to grow, appropriate software must be researched and implemented (Kelton, 1994).

B.   WHAT IS NEXT FOR SIMULATION

As proposed in the analysis of the ACM DL data, the simulation community may be poised for breakthroughs that no longer rely on processing speed and user-friendly software. The technological ceiling that Moore has predicted to be right around the corner may still hold true for the physical limits of computers and speed, but the future is wide open for simulation. At least it will be when improved analysis approaches, such as simulation optimization and efficient design of experiments, are readily available to the simulation analyst and in widespread use.

A recent leap of note was the June 2008 unveiling of a supercomputer known as "Roadrunner," a bank of machines capable of performing a quadrillion operations per second (petaflop). Such capability is so difficult to conceptualize that it is not unreasonable to say that it is approaching the limits of physics. It is hard to imagine a model that can act more efficiently than one that is being executed on the type of technology that can support petaflop-speed calculation. It would seem that such massive computing power would be enough to explore the limits of even the most complex model. However, even this level of computational power is not sufficient to provide a "brute-force" analysis for conducting thorough sensitivity analyses of the high-dimensional, complex simulation models that are pervasive in operations research, industry, business, and military simulations. In fact, a "brute-force" approach involving a single replication of a two-level experiment that involves 100 factors (examined at all 2^{100} potential combinations) would take over 40 million years for a simulation consisting of a single elementary operation (Sanchez & Wan, 2009). The moral is that advances in simulation cannot rely on advances in technology alone.

Fortunately, we need not wait for suitable analysis approaches to be developed. If simulation has started to level out in terms of technology-based advancements, then the design of experiment (DOE) methods have the potential to break the ceiling. Imagine the aforementioned two-level, 100 factor experiment as the "ceiling" for a machine as

capable as the Roadrunner. This is a stretch of the imagination as not even the most patient analyst can spare over 40 million years. Nonetheless, smart experimental designs can allow an analyst to identify the main effects of these 100 factors using as little as a few hundred carefully-specified combinations. For example, thanks to the DOE method of fractional factorials, commonly known as "screening designs" for eliminating unimportant factors from an experiment, there exists a "short program" for generating two-level experiments consisting of up to 120 factors that allows for all main effects and two-way interactions to be fit; the program develops a design "in under a minute," and this can easily be expanded into a design permitting the estimation of full second-order metamodels (Sanchez & Sanchez, 2005). Other useful designs include variants of nearly orthogonal Latin hypercubes (NOLH), and other space-filling designs. Similarly, advances in simulation optimization are providing new opportunities for simulation analysts.

Imagine a simulation model running on a computer with processing speed that allowed for each design point to take only one second to run. Such a scenario is feasible considering the processing speed of commercially available systems. Suppose that this model is comprised of 29 factors. The "brute-force" simulation approach would require over 17 years using a 2^{29} full factorial design. This takes the wind out of the ability to simply ride the technological wave. Apply the benefits of Latin Hypercube sampling and the same model would take under five minutes using an NOLH design. In other words, one can squeeze much interesting information out of a 29 factor experiment in less than five minutes (Sanchez, et. al., 2008). Efficient design also allows for flexibility in modeling, which is particularly useful in a world where the analyst must account for last-minute "good ideas" offered by his or her clients.

The theories behind DOE cannot take full credit for such flexibility. It is a combination of the computational capability available on modern computers and the "art" of experimental design. This marriage of the "brute" processing speed of today's machines and the power provided by efficient experimental design is an example of both sides of the simulation argument playing nice. The idea of optimality may very well exist under the hood of NOLH designs in that the combination of factor settings provides a

very precise estimation of ground truth. Furthermore, the concept of "optimality" permeates the practical real world where "sampling is expensive—the goal is to take no more samples than absolutely necessary." That is, even if you could spare the time to run as many replications of your complex model as you wanted, you do not have to (Sanchez et al, 2008). The optimality of efficient design is holding hands with the impressive capabilities of modern computers to create simulation models that represent a powerful force that is anything but "brute."

In short, the notional S-shaped curve of Figure 34 shows a limit that is reached once the simulation modeling and visualization capabilities become sufficiently advance and user-friendly. One can argue that this has already largely happened. Yet secondary S-curves that move beyond the modeling limitations are possible as we expand and institutionalize the simulation analysis capabilities. Capabilities that allow decision makers to easily move well beyond analyses geared toward examining a single system will be particularly valuable.

C.   RECOMMENDATIONS FOR FURTHER RESEARCH

While this research has explored technological advancements that have made significant impacts on simulation analysis, it concludes with a brief discussion on DOE and the academic pursuit of efficient design. These conclusions could inspire a future study into models that have taken particular advantage of such eloquent modeling techniques. Were a future study able to show a continual rise in simulation use, perhaps an argument would be found that could bridge the gap that divides analytical and simulation approaches to problem solving.

D.   CONCLUDING REMARKS

The exponential nature of simulation's growth over time has been made apparent without a need for deep exploration or analysis. The model fitting and analysis of influential years has pointed out some possible candidates for the most significant advances in the computing *and* academic fields. The takeaway is that simulation did not simply ride the technological wave and only take advantage of improved processing speeds. There have been serious contributions in terms of elegant thought, such as new

algorithms for pseudo-random number generation, simulation optimization, and efficient design. The impact of these contributions and of this research in general was confirmed by an independent correspondence between the researcher and Dr. David Kelton that took place only a few weeks prior to publication. Dr. Kelton provided several articles and insights that verified many of the significant simulation milestones and academic advancements in the subject of simulation that were mentioned in this research. As we educate the next generation of simulation practitioners and researchers, it is important to let them know that the "brute-force" stigma that has been so aggressively applied to simulation in the past is not an accurate reflection of the field. This research shows that many individuals in the OR community consider simulation to be the "first resort" when it comes to solving complex, real-world problems.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Birge, J. R. & Pollock, S.M. (1989). "Using parallel iteration for approximate analysis of a multiple server queueing system." *Operations Research*, *37*(5):769–779.

Bollen, K. A. & Jackman, R. W. (1990). "Regression diagnostics: an expository treatment of outliers and influential cases." In *Modern methods of data analysis,* Edited by J. Fox and J. S. Long, 257–91. Newbury Park, CA: Sage.

Cook, R. D. (1977). "Detection of influential observations in linear regression." *Technometrics, 19*(1), 15–18.

Conway, R.W. (1963). "Tactical problems in digital simulation." *Management Science*, *10*(1), 47–61.

Craddock, J. M. & Farmer S. A. (1971). "Two robust methods of random number generation." *Journal of the Royal Statistical Society.* Series D (The Statistician), *20*(3), 55 –56.

Fisher, R. A. (1925). "Applications of 'Student's' distribution." *Metron*, *5*:90–104.

Fritz, W. B. (1994). "ENIAC-a problem solver." *IEEE Annals of the History of Computing*, *16*(1), 25–45.

Geisler, M. A. (1959). "The simulation of a large-scale military activity." *Management Science, 5*(4), 359–368.

Harling, J. (1958). "Simulation techniques in operations research." *Operations Research, 9*(1), 9–21.

Hollocks, B. W. (2006). "Forty years of discrete-event simulation—a personal reflection." *Journal of the Operational Research Society, 57*:1383–1399.

Hurd, C. C. (1954). "Simulation by computation as an operations research tool." *Journal of the Operations Research Society of America, 2*(2), 205–207.

Kelton, D. W. (1994). "Perspectives on simulation research and practice." *University of Minnesota Supercomputer Institute Research Report,* 1–24.

Kottas, J. F. & Lau, H. (1982). "A four-moments alternative to simulation for a class of stochastic management models." *Management Science*, *28*(7), 749–758.

L'Ecuyer, P., Giroux, N. & Glynn, P. W. (1994). "Stochastic optimization by simulation: numerical experiments with the M/M/1 queue in steady-state." *Management Science*, *40*(10), 1245–1261.

L'Ecuyer, P., Giroux, N. & Glynn, P. W. (1994). "Stochastic optimization by simulation: convergence proofs for the G/G/1 queue in steady-state." *Management Science*, *40*(11), 1562–1578.

Lewis, P. A. W., Goodman, A. S. & Miller, J. M. (1969). "A pseudo-random number generator for the System/360." *IBM Systems Journal*, *8*(2), 136–146.

Lin, Y. (1993). "Commentary–Will parallel simulation research survive." *INFORMS Journal of Computing, 5*(3), 236–238.

Nance, R. E. & Sargent, R. G. (2002). "Perspectives on the evolution of simulation." *Operations Research, 50*(1), 161–172.

Parlar, M. & Sharafali, M. (2008). "Dynamic allocation of airline check-in counters: a queuing optimization approach." *Management Science*, *54*(8), 1410–1424.

Petersen, C. C. (1969). "GASP II simulation of parts for electronic fuze production." Scottsdale, AZ: Motorola Inc., Government Electronics Division.

Powers, M. J., Sanchez, S. M. & Lucas, T. W. (Forthcoming). "The exponential expansion of simulation in Research." *Proceedings of the 2012 Winter Simulation Conference*. Edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Raskin, J. R. (1970). "A tutorial on random number generation." *Computers and the Humanities, 4*(5), 307–311.

Rich, R. P. (1955). "Simulation as an aid in model building." *Journal of the Operations Research Society of America, 3*(1), 15–19.

Sanchez, S. M. (2008). "Better than a petaflop: the power of efficient experimental design." In *Proceedings of the 2008 Winter Simulation Conference*, Edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson and J. W. Fowler, 73–84. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Sanchez, S. M., Lucas, T. W., Sanchez, P. J., Nannini, C. J. & Wan, H. (2008). "Designs for large scale simulation experiments, with applications to defense and homeland security." In *Design of analysis and experiments, volume 3,* Edited by K. Hinkelmann and O. Kempthorne, Chapter 12. Hoboken, NJ: John Wiley and Sons, Inc.

Sanchez, S. M. & Wan, H. (2009). "Better than a petaflop: the power of efficient experimental design." In *Proceedings of the 2009 Winter Simulation Conference*, Edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 60–74. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Saran, C. (2005). "Time is running out for Moore's law as chip developments reach atomic levels." *Computer Weekly, 19*:12.

Schruben, L. W. (2011). Letter nominating Peter Lewis for the INFORMS Simulation Society's Lifetime Professional Achievement Award.

Teichroew, D. (1965). "A history of distribution sampling prior to the era of the computer and its relevance to simulation." *Journal of the American Statistical Association, 60*(309), 27–49.

Tocher, K. D. (1965). "Review of simulation languages." *Operations Research Quarterly*, *16*(2), 189–217.

Veith, T.L., Kobza, J. E. & Koelling, C. P. (1999). "Netsim: Java-based simulation for the world wide web." *Computers & Operations Research*, *26*, 607–621.

Wagner, H. (1969). *Principles of operations research with applications to management decisions*. Upper Saddle River, NJ:Prentice-Hall.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Professor Susan M. Sanchez
   Naval Postgraduate School
   Monterey, California

4. Professor Thomas W. Lucas
   Naval Postgraduate School
   Monterey, California

5. Dr. W. David Kelton
   Department of Operations and Business Analytics
   University of Cincinnati
   Cincinnati, Ohio

6. Professor Lee W. Schruben
   University of California, Berkeley
   Berkeley, California

7. Dr. Averill M. Law
   Averill M. Law & Associates, Inc.
   Tucson, Arizona

8. LCDR Matt Powers
   Naval Postgraduate School
   Monterey, California

9. Navy Modeling and Simulation Office
   Tucson, Arizona

10. Modeling and Simulation Coordination Office
    Alexandria, Virginia